

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 13, 2000, 23:12:37 ; Search time 2502.06 Seconds
(without alignments)
1991.508 Million cell updates/sec

Title: US-08-753-750A-1
Perfect score: 2792
Sequence: 1 atgataatgaatatcatca.....cattgaaatgaagttag 2792

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 972840 seqs, 892348106 residues

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:.*
1: gb_bal:.*
2: gb_bal:.*
3: gb_om:.*
4: gb_om:.*
5: gb_pat:.*
6: gb_ph:.*
7: gb_pl1:.*
8: gb_pl2:.*
9: gb_pr1:.*
10: gb_pr2:.*
11: gb_pr3:.*
12: gb_ro:.*
13: gb_sts:.*
14: gb_sy:.*
15: gb_un:.*
16: em_fun:.*
17: em_hum1:.*
18: em_hum2:.*
19: em_in:.*
20: em_om:.*
21: em_or:.*
22: em_ov:.*
23: em_pat:.*
24: em_ph:.*
25: em_pi:.*
26: em_ro:.*
27: em_sts:.*
28: em_sy:.*
29: em_un:.*
30: em_vi:.*
31: gb_htg1:.*
32: gb_htg2:.*
33: gb_in1:.*
34: gb_in2:.*
35: em_bal:.*
36: em_ba2:.*
37: em_hum3:.*
38: em_hum4:.*
39: gb_pr4:.*
40: gb_htg3:.*
41: gb_htg4:.*
42: gb_htg5:.*
43: gb_htg6:.*

44: gb_htg7:.*
45: em_htg1:.*
46: em_htg2:.*
47: em_htg3:.*
48: em_hum5:.*
49: gb_pl3:.*
50: gb_pr5:.*
51: gb_htg8:.*
52: gb_htg9:.*
53: gb_htg10:.*
54: gb_htg11:.*
55: gb_htg12:.*
56: gb_htg13:.*
57: gb_htg14:.*
58: gb_in3:.*
59: gb_htg15:.*
60: gb_htg16:.*
61: gb_htg17:.*
62: em_htg4:.*
63: em_htg5:.*
64: em_htg6:.*
65: em_htg7:.*
66: em_hum6:.*
67: gb_htg18:.*
68: gb_htg19:.*
69: gb_htg20:.*
70: gb_htg21:.*
71: gb_htg22:.*
72: gb_htg23:.*
73: gb_htg24:.*
74: gb_htg25:.*
75: gb_htg26:.*
76: gb_htg27:.*
77: gb_htg28:.*
78: gb_htg29:.*
79: gb_htg30:.*
80: gb_htg31:.*
81: gb_vil:.*
82: gb_vil2:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	% Match	Query Length	DB ID	Description
1	2781	99.6	6378	2	PHU73302	U73302 Pasteurella
2	531	19.0	3541	2	APU16019	U16019 Actinobacil
3	529.4	19.0	2826	5	A56682	A56682 Sequence 1
4	529.4	19.0	2833	1	APTBPI	249708 A.pleuropne
5	452.2	16.2	3447	1	APDNATFBB	254191 A.pleuropne
6	450.6	16.1	5083	2	APU16017	U16017 Actinobacil
C 7	411	14.7	5200	2	APL17915	Y17915 Actinobacil
C 8	345.2	12.4	10205	2	U32780	U32780 Haemophilus
9	340.8	12.2	5099	2	H1U15053	U15053 Haemophilus
10	340.8	12.2	5099	5	I79990	I79990 Sequence 4
11	332.6	11.9	5144	2	H1U15058	U15058 Haemophilus
12	332.6	11.9	5144	5	I79999	I79999 Sequence 10
13	332.4	11.9	4927	1	H1U10882	U10882 Haemophilus
14	331.2	11.9	5009	2	H1U15052	U15052 Haemophilus
15	331.2	11.9	5009	5	I79989	I79989 Sequence 3
16	331.2	11.9	5033	2	H1U15051	U15051 Haemophilus
17	331.2	11.9	5033	5	I79988	I79988 Sequence 2
18	314.4	11.3	4699	5	I79987	I79987 Sequence 1
19	271.8	9.7	3300	2	NGU16260	U16260 Neisseria 9
C 20	265.4	9.5	9955	2	AE002504	AE002504 Neisseria
21	262.4	9.4	3398	1	NMLBPAG	X79838 N.meningit
22	261.4	9.4	3171	1	NMIROA	X69214 N.meningit
23	261.4	9.4	5691	2	AF049349	AF049349 Neisseria
C 24	258.4	9.3	326301	2	NMA622491	AL162757 Neisseria

25	235.4	8.4	268	2	AF031694	AF031694 Pasteurel
26	231.4	8.3	2693	1	ACNTFFBAA	M85275 Actinobacill
27	231.4	8.3	2696	5	AR037085	AR037085 Sequence
28	231.4	8.3	2696	5	I11990	I11990 Sequence 1
29	231.4	8.3	2696	5	I21326	I21326 Sequence 1
30	215.6	7.7	8253	2	AF039315	AF039315 Moraxella
31	215.6	7.7	8266	2	AF039312	AF039312 Moraxella
32	193.8	6.9	2800	5	A36256	A36256 Sequence 3
33	193.8	6.9	4934	1	MNTBP12A	Z15129 N.meningit1
34	193.8	6.9	6114	2	AF058688	AF058688 Neisseria
35	190.2	6.8	8946	2	AF043131	AF043131 Moraxella
36	185.4	6.6	8937	2	AF043132	AF043132 Moraxella
37	181.2	6.5	3537	5	AR072081	AR072081 Sequence
38	166.2	6.0	29831	2	AF058689	AF058689 Neisseria
39	155	5.6	2809	5	A36258	A36258 Sequence 5
40	153	5.6	5472	1	MNTBP12B	Z15330 N.meningit1
41	150.4	5.4	11109	2	AE002402	AE002402 Neisseria
42	150.2	5.4	2762	1	MNTBP1B	X75166 N.meningit1
43	143.8	5.2	2916	1	NGOTBPB	M96731 Neisseria g
44	143.8	5.2	3286	5	AR072080	AR072080 Sequence
45	134.2	4.8	2653	1	MNTBPAG	X94533 N.meningit1

```

CDS
1309. .3063
/codon_start=1
/transl_table=11
/product="transferrin binding protein B"
/protein_id="AAB93472.1"
/db_xref="GI:27314433"
/translation="MFKLSQSVFLNLAALACSSNGGSFVDSQASAKVESQVOTQTPPKPPL
SQDDNSNARRTVSAGETBALLQPGFGSAKIPRRNLPGQKEDVAPIGDKIKETILR
PKGIVPEEYVGLGEPKADPSHTDRNKLVTDFNFVRSGYVHVHSGPKPKPEKILR
TGAPGVYVGLGEPKADPSHTDRNKLVTDFNFVRSGYVHVHSGPKPKPEKILR
SIDNVDDSEKPMWGTGETAFANKLTGLTVLRNIVYSRSEKQITTYDIDDAKLG
NRFSGSAKPKMILFKGSLKMSDTEGFGGEAQLAGKFLAADDKSVFYVFAFTRDAK
KDDSEAFADPFIKLLKMSDTEGFGNATHLIINNQLPIATAEAKDSFAEMKFDLL
TRTIDGVYVSCCNLDYVFKFGIYSEGNNSDTALQEXLVGERTALADLPCTGVYR
GTWGVYKSGSAGSVSNSESGTSLFVDVFNKKLNGKLIANDGVGEERPMPLTLE
GNLKGNGFGGTAKTNGSGNLDLPKSTNGTGVGHINTQFEGGFYGPRTATLGGIVQNT
TKDRVSIIFGGKRIQEK"
3151. .5943
/codon_start=1
/transl_table=11
/product="transferrin binding protein A"
/protein_id="AAB93473.1"
/db_xref="GI:27314434"
/translation="MTMKYHFRYSPVALTVLFALSHSYGAATENKKKIENNNDLAVLD
EVIYTSYHAFQRNEVTGLGVKYNHMSKNQILGIRDLTRYDPIGSIVVEOGRAS
SGVIRGVKRNVLSDVGLPOAHSYHTLGSANGGAINIEIYEYNTSRISLGSAGSA
EYSGAGHGAIGFRTKDAODILKEGOHWGLDSTSVASKNSHFLOSTAAAGEAGPEA
LVATHRHGEKTKIHSEANKLHNIRITGFENRYDFTQIPHRMLLEDLLLIYVEDCP
TLDTDPRARKLNRNDPVPPTPEYPERKKOLEQIPYTEQLSAQEYTGKORIANPY
LDKYSNFMFKGTFHFNSSHYLGALIEDTKTRYDIRDMQTPAYTTDDNLNLRNYY
EGDNLIDGLVFPRIQYGVSHVFKFDERHHKRLRGFTYKYPENNRWLDLSIKLSA
KODIELYSLRLHRSDDYPVDKNCRPTLDKWSMVTTERNNYQEKHVRVTHLEFDKAL
NAGQVFNQTHKLNLGLGDFDRNSLMDHGDMTAQYTKGYSYRGRGLNDPIYRRD
PRSETVLCNNITGDLNCEPKINGDSHFVSFDLVYSEYVDLGLGVKDFDGRHRRKS
DQPTLSYTRNWSWNGTILKPTFEVLSYRISNGFRVPAFYVGLGVKDFDGLKQNE
FYQARQSHQLEPEKSTHEIGVSFKQGYLDVYSFRNNYKMKIATACKRIISQKSV
FYVNRQIOVALNGILNVAKFDLGLTSLMPLDGFYSVYNNKRVKATCKRIISQKSV
NOPILDAIOPARYVLGVDYHPEKKGWIGITTTYSKAKNADEVAGTRHHGIGHRVDLGG
KLGTGSYTHDITGYINYNKYTLRGGIYVNTNRKSWESVRSQVNVNQRDGSNTYTR
FAGGRNFSLAPEMKF"
complement(6044. .6340)
/notes="similar to factor for inversion stimulation
protein"
/codon_start=1
/transl_table=11
/product="Fis"
/protein_id="AAB93474.1"
/db_xref="GI:27314435"
/translation="MLEQPTNPFLTVMNQAQOVNKPRLDRNVNAALKNYLSQLNG
EDPTLEYLVLSPIEHPLMDVMQYTRGNQTRATMLGINRGTLRKLLKLYKMG"

BASE COUNT 1977 a 1132 c 1382 g 1883 t 4 others
ORIGIN

Query Match 99.6%; Score 2781; DB 2; Length 6378;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2792; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 atgataagaatacatcatcttttcgcattatcacctgttgcccttaaacagtgatttgcct 60
|
|
|
Db 3151 ATGATAATGAAATATCATCTTTTCGCTATTTCACCTGTTCCTTAACAGTGTATTATTGCT 3210
|
|
|
Qy 61 cttctcatcattacacggtgctgcgactgaaaaataaaatacgaagaaaaataacagattca 120
|
|
|
Db 3211 CTTTCTCATTCACGCTGCTCGCACTGAAATAAAAAAATCGAAGAAAAATAACAGACTCA 3270
|
|
|
Qy 121 gctgtcttgatgaagtattgtgacagagagccattatgctctacgaacgtcacaacgaa 180
|
|
|
Db 3271 GCTGTCTTGATGAAGTATTGTGACAGAGAGCCATTATGCTACAGAACGTCACAAACGAA 3330
|
|
|
Qy 181 gtaactgcttggggaagtagtgaaaaattattcacgaaatgagtaaaatacaaatcttt 240
|
|
|
Db 3331 GTAACGTGCTTGGGAAGTAGTGAAAAATATCACGAATGACTTAAATCAAAATCTT 3390
|
|
|

```

Qy 241 ggatctcgtgatttaactcgtctatgacctggtgatttccgtggtggaacaagggtcgcggt 300
Db 3391 GGTATTCGTGATTTAACTCGCTATGACCTGTGATTTCCGTGGTGAACAAGGTGCGGT 3450
Qy 301 gcaagttagtgatgcatcgattcgaggttagataaaacccgtgctagcttacttgatgat 360
Db 3451 GCAAGTAGTGGCTATGCCATTGCGGCTAGATAAAAAACCGTGTGAGCTTACTTTGAT 3510
Qy 361 gggctaccacaagcgacacagttatcatcacgttagttcagatgctaaatgggtggtgcaatt 420
Db 3511 GGGCTACCACAGCGCAGATATCATACGCTAGGTTGAGATGCTAAATGGTGGTCAAT 3570
Qy 421 aatgagattgagatgaaacattcgttcaattgagtttaagcaaaaggagcaagttctcgc 480
Db 3571 AATGAGATTGATGAAACATTCGTTCAATTGAGTTAAGCAAAAGGAGCAAGTTCTCG 3630
Qy 481 gaatatgctctggtcgcatggtggtgctattggttttctgactaaagatgctcagat 540
Db 3631 GAATATGCTCTGGTGGCATGGTGGTCTATTTGTTTTCGTACTAAAGATGCGCAGAT 3690
Qy 541 attattaagagggcgacgattgggcttagatagtaagacctcttatgcccagcaaaaat 600
Db 3691 ATTATTAAGAGGGCAGCATTTGGGCTTAGATAGACCTCTTATGCCAGCAAAAAT 3750
Qy 601 agccattttttacagtctatcgcagcggtggtgaggggtggttttgaagcacttggt 660
Db 3751 AGCCATTTTACAGTCTATCAGCGCTGGTGAGCGGGTGGTTTGAAGCACTTGT 3810
Qy 661 attgcaactcaccgacacaggttaagagaccaaaatttcctcagagcaaaaataaa 720
Db 3811 ATTGCAACTACCAGACGGTAAAGAGACCAAAAATTCATCCGAGCAAAATAAATTA 3870
Qy 721 cataatattcggcgtataccggcttgaaatcgctacgactttaccacaaatccgcac 780
Db 3871 CATATATTCGGCGTATACCGGCTTTGAAATCGCTACGACTTTACCACAAATCCCGAC 3930
Qy 781 agaagtcctctggaggtactccttttaattgtggaagatacttggccacacttagattgt 840
Db 3931 AGAATGCTCTGGAGGATCTCCTTTAATTGTGGAAGATACTTGCCCAACATTAGATTGT 3990
Qy 841 actcctcgtgcaagggttaagtgaacgcgcgataaattcccagtgagaaactttccgaa 900
Db 3991 ACTCTCTGTGAAGGTTAAAGTTGAACCGCGATAATTTCCAGTGAGAACATTTCCGAA 4050
Qy 901 tatcgcctgaagcgcacacagcttagcagatcttctcagcactgagcagctctca 960
Db 4051 TATACGCTGAAGACGCAACAGCTTGAGCAGATTCCTTATCGCAGCTGAGCAGCTCTCA 4110
Qy 961 gcccaagaatataccggttaaagatcgcttgcaccaaaacctttagattacaagagtaat 1020
Db 4111 GCCCAAGAATATACGGTAAAGATCGCATTCACCAAAACCCCTTTAGATTACAAGATAT 4170
Qy 1021 tctgtttttatgaagtgttgctatcaactcactcactcactcactcactcactcactc 1080
Db 4171 TCTGTTTTTATGAAGTTTGGGTATCACCTCAACTCGTCTCATTTATCTTTGGCGCAATCTTA 4230
Qy 1081 gaagatacaaaaacgcctacgatatcgctaatcgctaatgcaaacgcagcttactatcaaaa 1140
Db 4231 GAAGATACAAAACAGCTAGCATTCGTTATCGTATGCAACGCCAGCTTACTATACAAA 4290
Qy 1141 gacgatattaacttactcacttaggaactatggtttatgaagggtgataatatttttagatggc 1200
Db 4291 GACGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4350
Qy 1201 tagtggttcaagcgaagatcccttatggttggttggttggttggttggttggttggttg 1260
Db 4351 TTAGTGTTCAGGCCAAGGATCCCTTATGGTGGCTATAGCCATGTGAAGTTTTTTCAT 4410
Qy 1261 gaacgtcaccacaacgcgtttaggattccctataataataaaccagagaaataatcgc 1320
Db 4411 GAACGTCACCACAACGCTGTTAGGATTCACCTATAATTAACACAGAGAAATAATCGC 4470

Qy 1321 tgggttgatagcattaaactcagtcggtgataaaacagataattgaactatatagccggcta 1380
Db 4471 TGGTTGGATAGCATTAACACTCAGTCGGATAACAAGATATTGAATATATAGCGGCTA 4530
Qy 1381 catcgttgcaatttagcgattatccttggttagataaaaaatttccgcccacacttgat 1440
Db 4531 CATCGCTTGCATTAGCGATTATCCTGTGTAGATAAAAAATTCGCCCGCACTTTGGAT 4590
Qy 1441 aaacttggtctatgctacgaactgagcgttaataattaccagaagaaacatcgtgctatt 1500
Db 4591 AAATCTTGGTCTATGATCGAATCTGAGCGTAAATTAATTAACCAAGAAACATCGTGTCAAT 4650
Qy 1501 catttagaatttgataaagcgcctaaatgctggttcaagcgtatttaacccaaacccacaaa 1560
Db 4651 CATTTAGAAATTTGATTAAGCGCTAAATGCTGTCAGGCGTATTTAACCAAAACCCACAAA 4710
Qy 1561 ctgaatttagggttgggctttagcgtatttaattcgcttattggtatcattggggatatgact 1620
Db 4711 CTGAATTTAGGTTGGGCTTTGATCGATTTAATTCGTTATGATCGGATATGACT 4770
Qy 1621 gcccaatataccaaagcggttataccagctaccgcggttagagcgtttagataaacc 1680
Db 4771 GCCCAATATATCAAAAGCGGTTATACAGCTTACCGCGGTAGAGGCGGTTAGATAATCCA 4830
Qy 1681 tataattatccgcgcgtatccacgcagattgaaacggttatcttctgtgtaataacacgc 1740
Db 4831 TATATTTATCCCGCATCCACGCGTATTGAAACGGTATCTTTGTGTAAATAATACACGC 4890
Qy 1741 ggcacatctttaaactgtgaacgcgcgttaaaattaaagcgtatgcccatttctgttagcttc 1800
Db 4891 GGCACATCTTTAAACTGTGAAACCGCGTAAAAATTAAGGCGTAGCCATTTTGTAGCTTC 4950
Qy 1801 cgcgatctagtgataaagcagatgctgatttggatttggatttgggttgcgtttttagacaacat 1860
Db 4951 CGCATCTAGTGTATAGCGAGTATGTGGATTGGGATTAGGGGTGCGGTTTTGTATCAACAT 5010
Qy 1861 cgattaaactgctgatccgtgacacttagccgaacttatcgaaaattgctcttggaaat 1920
Db 5011 CGATTTAAATCTGATGATCCGTGGACACTTAGCCGAATTTATCGAAATTTGCTTTGGAA 5070
Qy 1921 ggtggattacgcgttaaaccaacagagtttgtatcgcttttcttcttcgcgaatttcaaacggt 1980
Db 5071 GGTGGGATTACGCTTAAACCAACAGAGTTTGTATCGCTTTCTTATCGCATTTCAACCGT 5130
Qy 1981 tttagagctcctgcatctctatgaacttattgttaaacgtgataatttggctttaaagat 2040
Db 5131 TTTAGTGGCTGCTGATCTATGAACCTTTATGGTAAACGTGATCATATTTGGGCTTTAAAGAT 5190
Qy 2041 aacgaatgctgcaacgcgcgcacacttagccacagtttagagccagaaaaatcgcactaat 2100
Db 5191 AACGAATATGTCAACGCGCNAAGTAGCCAGTACCCAGTTAGAGCCAGAAAAATCCGCTAAT 5250
Qy 2101 catgagattgagtagctttaaagggttaatttggttacactttagtgagctatttcogt 2160
Db 5251 CATGAGATTGAGTTAGCTTTAAAGGTCATTTGGTTACCTTTGATGTGAGCTATTTCGGT 5310
Qy 2161 aataactataaaaatgattgcacagcatgtaaaagataataacaaaaatcacactgt 2220
Db 5311 AATAACTATAAAAATATGATTGGCAGCATGTAAAAGATAAATACAAAAATCACACTGT 5370
Qy 2221 ttctataactaccataatattcaagatgtagcactaaacgggataaaatttagtcgcctaaa 2280
Db 5371 TTCTATAACFACCATAATATTCAGATGTAGCCTAACGGGATTAATTTAGTCGCTAAA 5430
Qy 2281 tttagcttacacggtattttatctatgctgcagatggtttttttatcatcagttgcttat 2340
Db 5431 TTTGACTTACACGCTATTTTATCTATGCTGCCAGATGGTTTTTATTCATCAGTTGCTTAT 5490
Qy 2341 aaccgtgtaaaagtaaaagagcggaactaaccaactaacgactcagatagcgtaaacgat 2400
Db 5491 AACCGTGTAAAGTAAAGAGCGGAAACTAACCGACTCAAGACTCGATAGCGTAAACGAT 5550
Qy 2401 cctattctagatgctgattcagccagcagcgtatgtgcttggattcggctacgataccca 2460

Db 1376 ATAAAGATGCAGAGTCGTTCTCAGAAATATCACCCGAGTAGGAGTGGAACCAACAGCAGC 1435
Qy 749 aaatcgcctacgaacttaccacaaattccgcacagaatgctcctcctgagagatctccttttaa 808
Db 1436 TTGATAGCTCAATAGATATATCTCGCAGCAGGAATAATACACATACATTTAGGCTGGTTTT 1495
Qy 809 ttgfggaataacttgcacacattagattgtagtactcctcgtgcaagggttaagtgaacc 868
Db 1496 TGATTAAGATGAATGTCCAACGTTAGATTGTACGCCGAACACAGATGGCTAGGGTGACAA 1555
Qy 869 gcgataatttcccagtgagaaacttccggaatatatcgcctgaagagcgcgaacagcttg 928
Db 1556 AAGATAGCCCATCTTCGGTCTTACCTGAAATATATCTCTGAGGAAACAGAGCTTATG 1615
Qy 929 agcagattccttatcgcactgagcagctcctcagcccaagaatatcacccggttaagatcgca 988
Db 1616 AGAACCAAAAACATATTACAGAGCGTCTAAATGCTCAGGATTACACTGGTGAAATATAGAG 1675
Qy 989 ttgcacaaaccccttttagattacaagagtaattctgttttttttagaagtgttgctatcact 1048
Db 1676 CTTTACCTGATCCCGCTTAAATATAAATCTGATTCTTGCTGGTTTAAATTTAGGATACACAT 1735
Qy 1049 tcaactgctcattatcttgcgcacactcttagaagatacaaaaaacacgcctcacgatccc 1108
Db 1736 TCTCTCGAAACATATTGCTCCTGGTACTTTATGAACATAGCAAAACAGCGTTTACGACACC 1795
Qy 1109 gtgatatgcaaacgccagcttac----- 1131
Db 1796 GAGATATGACCTATACCGCTTATTGGCAACCATCGGATTTACTTTAGAACCTGGTAGAAATT 1855
Qy 1132 --tatacaaaagacgataattaaactatacacttaggaactatgtttatgaaggggataata 1189
Db 1856 GGTATCCAATGAATATGCTAAAGGATTTATATCGTGAATAATGCTTTTAGATGGTGTGCTA 1915
Qy 1190 ttttagatggcttagtgttcaagcgaaggtcccttatgggttgcctatagccatgtga 1249
Db 1916 TTGACTACTTTACCGAGATGGTGTGAATCATCAAAAGGTTTACGTTGGGCAAAAGCTC 1975
Qy 1250 agttttttgaaagcgcacacaaacgctgttttagattcaactataaataat---aaac 1306
Db 1976 GTTTTATTAGGAGTGGCACACTCGTGATCGCTTAGGTGCTTTATATCGTTTATACCAATC 2035
Qy 1307 cagagaataatcgtgttggatagcattaaactcaagtgcggtgagataaacaagattgaac 1366
Db 2036 AAGATGGAATTCGTGTATGATAGACTATCTTGGAGTTTGGATCAGCAAAAAATTAAT 2095
Qy 1367 tatatagcgcgtacatcgtctgcatgtagcattacctgtggttagataaaattgccc 1426
Db 2096 TATCTACCCGCTTGAGAGAGAACAACTGTTCGGAATATCCAAACCATAGTAAGAATTGCC 2155
Qy 1427 gcccgactttggataaaacttggctctatgtatcgaactgagcgtlaataattaccagaata 1486
Db 2156 GTGCAACTCTTGATAAACTTTGGCTTCAACTAAAATGACCAAAAGTTCTTTATGAAGAA 2215
Qy 1487 agcatcgtgctatcatttagaatttgataaagcgcgtaaatgctggtgtaagggcgtattta 1546
Db 2216 AACACGACACTATTACGCTCTCGTTAGATAAAACCGTACAAACGGGATTTGGG----- 2267
Qy 1547 accaaacccacaactgaatttaggggttggctttgctgcatgatttaattcgtttggtatc 1606
Db 2268 ----TAAACATCAATTAATATGTTATTAGGTTTCAGACCGTTTCAATTCACCTTAAAC 2323
Qy 1607 atgggggatgactgcccataatacacaagcgcggttataccagctaccgcggttagagggc 1666
Db 2324 GCCAGAAATTTTGGTGAATTTCTGTGGGAACCTTGGCATCGTATCAGAGGTAAACGGTT 2383
Qy 1667 gtttagataatcccatatatttatgcgc-----gcgatccacgcagattgtaaacgggtat 1720
Db 2384 ATAAAGATACACCTTACATCTATGAGCTTAAAGATCAGGCAATTTATAGTAAAAATGAAT 2443
Qy 1721 ctttgtgtaataatacacgcggcgacatctttaaactgtgaaccgcgttaaaattaaagcgc 1780
Db 2444 GTGATTATAGTGGCACTATTTCAGGTAGGGCTGATTGTGCTTACAAGTAAAAATCAAGGGC 2503

Qy 1781 atagccattttgttagcttccgcg----atctagtataaagcagagtatgtggtattgggat 1837
Db 2504 ATATCACTACTACGCTCTGAGAGATAAATTTGCCATAACCAAGTATTTGATGATTGGTT 2563
Qy 1838 taggggtgcgttttgatcaacatcgatttaaatctgatgctcgttggaacttagccgaa 1897
Db 2564 TGGGTTACCGTTTCGATAAGCAATAAATTCGCTAGCACTCATCGTGGGCAAAACAAGCG 2623
Qy 1898 ctatcgaaattggcttgggaattgggtggtattacgcgttaaaccaacagaggttggatcgc 1957
Db 2624 ATTATAAAACAGTCGCGTGAATATTGGCATAGTCGCAAAACCAACGCTCATCTCTATCGC 2683
Qy 1958 ttcttctatcgacttcaaacgcgttttagagtgctcgtcattcttatgaactttatggttaac 2017
Db 2684 TCTCTTATCGAGCATCATCTGGCTTTAGAGTGCCAAAGTTTCCAAAGAGTATTTGGCTTAC 2743
Qy 2018 gtatcatattgggcttaaaagataaagaaatatgtgcaacgcgcgcaacgttagccacag 2077
Db 2744 GTTATGATGTGCAATGAAAGGCTCCAGCGATGCTTA-----CCAAAAACAGAGAAGT 2797
Qy 2078 tagagccagaaaaaactgactaatcatcatgagattgaggttagctttaaagggtcaatttggtt 2137
Db 2798 TATCTCTGAAAAATCCTTAACCAAGAGGTTGCTGGCACTTTCAAAAGGTGATTTTGGTG 2857
Qy 2138 accttgatgtgactatttcccgtaactataaactataaaaaatgattgacgacgcatgtaaaa 2197
Db 2858 TCGTTGAAGTCAGTTATTTCAAAAAATGACTATAAGCAGTTAATTGCTCCAGCAGA---- 2912
Qy 2198 gaataatacaaaaaactcacactgtttctataactaccataataattcaagaatgtagcaactaa 2257
Db 2913 -AAGATGTCATCAAACTCAATCAATGATTAACTATTATTAATGTGCAAGATATTAAATTTGG 2971
Qy 2258 acgggataaaatttagtcgctaaatttgaacttacacggtattttatctatcgtgcagatg 2317
Db 2972 ACGCATTAATCTATTGGTTAAGCTAGATTGGAATGGGGTATTGTGATAAATTCCTGAGG 3031
Qy 2318 gtttttattcatcagttgctttaaacgcgttaaaagtaaaagcgcggaagcgaactaacccgact 2377
Db 3032 GCATTTACACAACATTTGGCTTATAGCAAAATGCGAGTAAAGAGGTGAA---AAACTATC 3088
Qy 2378 caagactcgtatagcgttaaacgactctatttagatggattcagccagcagcgttatg 2437
Db 3089 AAGGGTATATGAATATTTCGTTCTCCATTGTTAGATACCATTCAGCCCTGCTCGCTATGTTG 3148
Qy 2438 ttggtatcgcgtacgatacccccagaagaaaaatgggaattggcattactaccacctatt 2497
Db 3149 TAGGAGTGGGTACGATCAGCCAGATGAAAAATGGGCGTGAAATCTAACAAATGACACACT 3208
Qy 2498 ctaaaagccaaaacgcgcgatgaggtggcagcgcacgctcatcagcggtatagatcgcttga 2557
Db 3209 CCAGTGGAAAAAATCCAAATGAGTTAAGAGGTAATGAACAAGTCGGTTTTGCCAAATTATG 3268
Qy 2558 ttttagtgcgc-----aaactgacgcgttcttggtcacaccatgatatcaccggttaca 2610
Db 3269 AGCGAAGTGCACGAAGAAAAAACACGCTTCTTGGCATACCTTTGACTTAAACGGGATATA 3328
Qy 2611 tca---attataaaactcacacctacggtgaggaatttataatgtgactaaactgtaaat 2667
Db 3329 TCACCCCTTGGAAACATPACAAACGGTACGAGTGGCTGATATATACCTGATGAATTTATCGTT 3388
Qy 2668 attcatttgggaatcagtcgcgaatcccggttgtagtcagtaaaacacagacgcgg---g 2724
Db 3389 ACACCACTTGGGAATCCGTACGTCATCTTCGCTTATGCAATTCATCAGCATACTAACG 3448
Qy 2725 gtagcaattacactcgtattggcgcctccggggggaatttcaagtttagcattgaaatga 2784
Db 3449 TAAAAGACTATGCAAGGTATGCGCCGCGTGAATAATTATGTTTTCATTCGAAATGA 3508
Qy 2785 agtttta 2791
Db 3509 AATTCTA 3515


```
QY 1838 taggggtgcgttttgatcaaacatcgatttaataatctgatgcgtggagcaacttaagccgaa 1897
Db 1854 TGGGTTACGTTTCGATAGCAATAATTCGTPAGCACATCGCTGGCGCAATCAAGCGG 1913
QY 1898 ctatcgaattggtcttggaatggtggattacgcttaaaccaacagagtttgtatcgc 1957
Db 1914 ATTATAAACACAGTGGTGGATATATGGCATAGTCGCAAAACCAACGTCATTCCTATCGC 1973
QY 1958 ttcttatcgcatcttcaaacggttttagagtcctgcattcttatgaacttttatgtaaac 2017
Db 1974 TCTCTATCGAGCATCATCTGCTTTTAGAGTGCCAAAGTTTCCAAGAGCTATTTGGCTTAC 2033
QY 2018 gtagcatattgggtttaagaatacgaatatgtgcaacgcgcgcaacgtagccaccagt 2077
Db 2034 GTATATGATGGTCAATGAAGAGCTCCAGCATGCTTA-----CCAAAAACAGAGAAAT 2087
QY 2078 tagagccagaaaaatcgataatcatgatgagtgaggttagcttttaagagtcgaatttgtt 2137
Db 2088 TATCTCTGAAAAATCCTTAACCAAGAGGTGCTGCGACTTTCAAGGTGATTTTGGTG 2147
QY 2138 accttgatgtgagctatttcccgtaaacataactataaaaaatgatgtagcgacagcatgaaaa 2197
Db 2148 TCGTTGAAGTCAGTTATTTCATAAATGACTATAAGCAGTTAATTGCTCCAGCAGA---- 2202
QY 2198 gaataatacaaaaaatcaactgtttcttataactacataataattcaagatgtagcaactaa 2257
Db 2203 -AAGAAATGCATCAAACTCAATCAATGATTAACATATTTAATGTGCAAGATATAAATGG 2261
QY 2258 acgggataaatttagctgataattgacttacacggattttatctatctgctgccagatg 2317
Db 2262 ACGGCATTAACTTATGTAAGCTAGATATGGAATGGGGTATTTGATAAAATTCCTGAGG 2321
QY 2318 gttttattcatcagttgcttataaccggtgtaaaagataaaagagcggaactaacccgact 2377
Db 2322 GCATTTACACACATGCTTATACCAATGCGAGTAAGAGGTGAA---AAACTATC 2378
QY 2378 caagactcgatgcgttaaacgcatctctatctagatgcgattcagccagcagctatgctgc 2437
Db 2379 AAGGGTATATGAATATTCGTTCTCCATGTTAGATACCATTCAGCTGCTGCTATGTTG 2438
QY 2438 ttggattcggctacgatcacccagagaagaaatgggaattgggaattactactaccactatt 2497
Db 2439 TAGGAGTGGGTACGATCAGCAGATGAAGAAATGGGGGTGAATCTAACAAATGACACACT 2498
QY 2498 ctaaagccaaacacccgcatgaggtgagcgacacgcgtcaccaggtacacatcgcttga 2557
Db 2499 CCACTGGAAAAATCAATGAGTTAAGAGGTAATGAACAACGTCGTTTGCCCAATTATG 2558
QY 2558 tttagtggtgc-----aaactgacgggttcttggtacacccatgatattaccgggtaca 2610
Db 2559 AGCGAACTGCCACGAAGAAAGAACACATCTCTTGGCATACCTTTGACTTAACGGGATATA 2618
QY 2611 tca----attataaaactacacttacgtgaggaatttataatgtgactaatcgtaaat 2667
Db 2619 TCACCCCTTGGAAACATACACGGTACGAGTGGCGGTATATAACCTGATGAATTATCGTT 2678
QY 2668 attccacttgggaatcagtcgcccatacgggtgtgtaagtcagtaaaccaagaccgg---g 2724
Db 2679 ACACCACTTGGGAATCCGTACGTCATCTTCGCTTAATGCAATTCATCAGCATACTAAG 2738
QY 2725 gtagcaattacactgattggcgtccggggagagaatttcagtttagcatttgaatatga 2784
Db 2739 TAAAGACTATGCAAGGTATGACGCCGCCGGTGAAGAAATATATGTTTTCATTTCGAATGA 2798
QY 2785 agtttta 2791
Db 2799 AATTCTA 2805
```

RESULT 5

APDNATFBB

LOCUS

DEFINITION

A.pleuropneumoniae

tfbb gene

encoding

transferrin

receptor.

20-SEP-1995

BCT

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

254191
254191.1 GI:994790
tfbb gene; transferrin receptor.
Actinobacillus pleuropneumoniae.
Actinobacillus pleuropneumoniae.
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Actinobacillus.
1 (bases 180 to 2873)
Wilke,M.
Thesis (1995) Institut fuer Mikrobiologie, Tieraerztliche
Hochschule Hannover
2 (bases 1 to 3447)
Wilke,M.
Direct Submission
Submitted (19-SEP-1995) Wilke M., Tieraerztliche Hochschule
Hannover, Institut fuer Mikrobiologie, Bischofsholer Damm 15,
Hannover, Germany, 30173
Location/Qualifiers
1. 3447
/organism="Actinobacillus pleuropneumoniae"
/plasmid="pTF205/023"
/strain="AP205"
/db_xref="taxon:715"
/clone_lib="lambda 2001-library"
/notes="serotype:7"
180..2876
/gene="tfbb"
180..2876
/gene="tfbb"
/function="binding of porcine transferrin"
/codon_start=1
/transl_table=11
/evidence="experimental"
/product="transferrin Receptor"
/protein_id="CAA90896.1"
/db_xref="GI:994791"
/transl_table="MKNKLNLSLISLFAVSQAQVQLNDVYVTGKKAHKKE
NEVTGLKVVKTPTDLSKSEQLVGRDLTRYDPGISVVEQGRGATGYSIRGVDNRVG
LAIDLQIOISVYVSSSSGAINETEVENLRISIOISKASSSEFGSLGSGVQPR
TKEVSDIIKFGQWGLDTKSAYSSKNOUWNSLAFTHNGFESLVYIIRHDKRETKA
HKDAESKSIQRVDLADNPQSNFVKVNDPCSLNCEPQQVGVSDYDQKITEQL
SAKEYTEERALPDVYKSDLSWLKGLYSLPKYHVGYSKGLRWAKRFIDENH
WQSLDLKNGDKWPMNNAKGLYRDNALDGVADYFTEDGVKSSKGLRWAKRFIDENH
TORIALGYRYTNOGDNRLDLRLSLFDOOKINLRLRENNCSEYPTIDNCRATLD
KLWSPKNEQSSVEERHDTIQLSDKTQVGLCKHQLMELLGSDRFNSTLKRIHELSE
ALRDNFAITKIDIGLYRFDKHKFRSHRANQGDYKNSAWNIGIVAKPTLSLSY
RASSGFRVPSQELFLGRYDGMKSSDAYQKTEKLSPEKSLNQEAATKFGDFGVY
VSYFKNDYKOLIAPAERHQTQSMINYNFVQDIKLDGINLGLDWNQVDFKIPGTY
TTLAYSKMRVKEVKGYNIRSPLLDTIOPARYVGVGVYDQDEKQWYNLMTHTSS
GKNPNELRGNEQVGFANYERTAKPKTRSHHTFDLGYITPWHKTTVRAGYNNLMYR
YTTWESVRQSSLNALHOHTNVKDYATRAHPRNVVSEFMKF"

BASE COUNT
ORIGIN

1160 a 610 c 719 g 958 t

Query Match 16.2%; Score 452.2; DB 1; Length 3447;
Best Local Similarity 52.4%; Pred. No. 7.8e-96;
Matches 1393; Conservative 0; Mismatches 1163; Indels 103; Gaps 14;

QY 158 atgtctcacgaacgtcaaaacgaagttaactggtctgggaagtagtgaataatcacg 217

Db 295 AAGCACATAAAAGAGAGNACGAGTACGAGGCTTAGGGAAGTAGTGAACACACAGATT 354

QY 218 aaatgagtaaaatcaaatctcttggtattctgatttaactcgctatgacctggtattt 277

Db 355 CTCCTAGTAGAGGACGAGTGTAGGAATCGAGATCTGACTCGCTACGATCCGGTATT 414

QY 278 cgggtggtgaacaaggtcggtgcaagtagtgcttatgccattcgattgaggtgataaaaa 337

Db 415 CTGTAGTAGCAAGGAGGAGGTCAACGACAGGCTACTCAATTCTGTGGGTAGATCGTA 474

Qy	338	acsgtgtcagcttacttctgtatggtgctaccacaaagcgcagattatcatcactcaggtt	397
Db	475	ATCGTGTGGGCTTGGCATTAGACGGTTTGGCACAGATTCAATCTCTTAAGTCAATATT	534
Qy	398	cagatgctaagtgtgtgcaataataatgagattgattgaaaaacattcgtttcaattgagt	457
Db	535	CACGTTCCCTCAAGCGGTGCATTAATGAATAAGAAATACGNAAAATCTCGGTTCGATCCAAA	594
Qy	458	taagcaaaagagcaagcttctgcggaatatggctcgtggtgcacaggtggtgctatttggtt	517
Db	595	TTAGTAAGAGGAGCTAGTCTCTCTGAGTTTGGCAGTGGCTCGCTAGGCGGTTCGGTGC AAT	654
Qy	518	ttcgtactaaagatgcgcaggatattataaagaggggcagcattggggcttagatagta	577
Db	655	TCCGTACCAAAAGGTAAGCGACATTATTAAAGCCAGGCAATCTTGGGACTAGATACCA	714
Qy	578	agacctcttatgccagcaaaatagccattttttacagctctatcgcagcggctggtagg	637
Db	715	AAAGTGCCTACAGCAGCAAAAATCAACAATGGTTAAACTCACTTTCCTTTTCGGGGTACT	774
Qy	638	cgggttggtttgaagcacttgattatgcaactcaccgcacggttaaagagacacaaattc	697
Db	775	ACAAATGGCTTTGAGTCTCTGTGATTACACTCACCGTGATGGTAAGGAACAGAAAGCTC	834
Qy	698	attccagggcacaaataattaaacataatattcggcgtataaccggctctgaaaaatcgct	757
Db	835	ATAAGGATGCAGAAAGCCGTTCTAAGAGTATTTCAGAGATGGATCTCGCAGACAAT---	890
Qy	758	acgactttacccaaattccgcacagaatcgtctcgtgagga tctccttttaattgtggaag	817
Db	891	-----AATCCTCAAGGTAGCAATTTGGTTTAAAGTTAAGAA	925
Qy	818	atacttgcccaacattagattgtactctctgcgaaggtttaaqttagaacccgcgaatt	877
Db	926	TGATTGTCC-----AAGTTTGAACATGCGAGCCAGCCAGCAGGTGTGGAGTCTTATGA	979
Qy	878	tcccagtgagaacatttcgggaataacgcctgaagacgcgaacacagctttgagcagattc	937
Db	980	TGGTCAGAGTATTATACAGAGCAACTCAGTGCTTAAAGAATATACAGGTGAAGAACGAGC	1039
Qy	938	cttatgcgactgtagcagctctcagcccagaata taccggttaaagatcgcattgcaccaa	997
Db	1040	CTTGCC---CTGATCCAGTGAATAATAATCTGATCTTGGCTGGTAAATAGGTATATA	1095
Qy	998	accctttagattacaagataattctgtttttatgaagtttggctatcaccttcaactcgt	1057
Db	1096	GCCTTTCCACCAAAACATTATGTGGCTGGTGGTTTATGACATAGCAAGCAACGCTATGATA	1155
Qy	1058	ctcatctattggtgcgaatcttagaagatacaciaaacgcgtcacgatatccgtgatatgc	1117
Db	1156	CTCGGATATGACCTATCTCGCTTATTGGCAGCTATCAGATTTAAAAAATGGTGATAG-	1214
Qy	1118	aabscgcagcttactatacaaaagacgataattaactttacacttaggaactatgtttatg	1177
Db	1215	-----TGGTATCCAATGAATAATAGCTTAAAGGATTATATCGTGATAAATGCTTTAG	1263
Qy	1178	aagggaataatttttagatggccttagtgttcaagcccaaggatcccccttatgggttgcct	1237
Db	1264	ATGGTGTGTCTTTTGACTACTTTACGGGAAGATGGTGTGAAATCATCAAAAGGTTTACGTT	1323
Qy	1238	atagccatgtgaagttttttttagtagacgttcacccacaaacgcttggttaggattccacctata	1297
Db	1324	GGGCAAAAGCTCTTTTATTGACGAGTGGCAGACTCGTGATCGCTTAGGTCTTTATATC	1383
Qy	1298	aatat---aaaccagagaataatcgtcgtgtgtgtagcatctaaactcagtcggtggaataac	1354
Db	1384	GTTATACCAATCAAGATGGAATAATCGTCTGATTGTATAGATAGACTATCTCTTGAGTTTCGATCAAC	1443
Qy	1355	aagatatgaactatatagccggctacatcgcgttcgattgtagggattatcctcgttgtag	1414
Db	1444	AAAAAATTAATTTATCTACCCGCTTGAGAGAGAAACAACCTGTTCCGAATATCCAACCTATG	1503
Qy	1415	ataaaaattgcccgcgcagcttttqataaaaatcttgctatgattcagctcagctcagcgaata	1474

Db 2557 CAATGACACACTCCAGTGGAAATAATCCAAATGAAGTTAAGAGGTAAATGAACAAGTCGGTT 2616
Qy 2546 acatgcgcttgattagtgagc-----aaactgaccggttcttggttacaccatgata 2598
Db 2617 TTGCCAATTATGAGCGAAGTCCACGAGAAAGAACACAGCTTCTTGGCATACCTTGA 2676
Qy 2599 ttaccggttaacata---attataaaactacacaccttacgttgaggaatttataatgtga 2655
Db 2677 TAACGGGATATATACCCCTTGGAAACATACAAACGTCAGGCTGGCGTATATAACCTGA 2736
Qy 2656 ctaactcgtaaatattccacttggaatcagtcgccaatccggttggaatcagtaaac 2715
Db 2737 TGAATATTCGTTACACCACTTGGGAATCCGACGTCATCTTCGCTTAAATGCAATTCAT 2796
Qy 2716 aagaccgg---ggtagcaattacacactgattggcgctccggggagagaatttcagtttag 2772
Db 2797 AGCATACTAACGTAAGACACTATGCAAGGTATGCAGCGCCCGGTAGATAATTTGTTT 2856
Qy 2773 catttgaataagagtttta 2791
Db 2857 CATTCGAATGAATTCCTA 2875
RESULT 6
APU16017 5083 bp DNA BCT 09-NOV-1995
LOCUS Actinobacillus pleuropneumoniae transferrin binding protein 2
DEFINITION (tbpB) and transferrin binding protein 1 (tbpA) genes, complete
cds.
ACCESSION U16017
VERSION U16017.1 GI:1055209
KEYWORDS Actinobacillus pleuropneumoniae.
SOURCE Actinobacillus pleuropneumoniae
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Actinobacillus.
REFERENCE 1 (bases 1 to 4575; 1 to 1855)
Gonzalez,G.C., Yu,R.H., Rostek,P.R. Jr. and Schryvers,A.B.
TITLE Sequence, genetic analysis, and expression of Actinobacillus
pleuropneumoniae transferrin receptor genes
JOURNAL Microbiology 141 (Pt 10), 2405-2416 (1995)
MEDLINE 96036198
REFERENCE 2 (bases 1 to 5083)
Schryvers,A.B.
AUTHORS Direct Submission
TITLE Submitted (17-OCT-1994) Anthony B. Schryvers, Microbiology and
JOURNAL Infectious Diseases, University of Calgary, 3330 Hospital Drive NW,
Calgary, Alberta, T2N 4N1, Canada
FEATURES
source Location/Qualifiers
1. 5083
/organism="Actinobacillus pleuropneumoniae"
/strain="H49"
/db_xref="taxon:715"
/clone="pBAP11"
/note="serotype 7"
197. .200
206. .1849
/gene="tbpB"
206. .1849
/gene="tbpB"
/codon_start=1
/transl_table=11
/product="transferrin binding protein 2"
/protein_id="AAC43484.1"
/db_xref="GI:1055210"
/translation="MHFKLPALFTSLFLVACSGGSGFDLEDVRPNKTTGYSKREE
YKDVETAKKEQELGEMEPALGVVVPVSSFNKKVVDISDIEVINGLDDVPYKA
NSSKYNYPDKTKDSSLOQYVRSVVIDGSHSGNKKVYVYKGNSPAKELPVNOLLY
TGSWDFSNANLNNEGRPNVINDYYTKFGRVGLVSGDAKPAKHKYSQEVDEFA
TKMTKLSKKEKTIYVNADIRGNRTGAATSDKNKGESINFFSADQSLEGGF
YGPRAEMAGKVFANDKSLFAVFSKINGNSVPTVRIIDASKIDLTFNFSLELNFGD
ASVLIDGRIKLAGSGFTNKHTEIINGKTMVAVACCNLEYKFKFGLWQAAGGKPE
NNSLFLQGETATDKMPKGGNYKYGTDWQVSKENNWNWATADDDRKAGYRTFEDVDF

GKNLSGKLFDRKGNVPFTVDPKIDKNGFTGAKTSDAGFALDSSSSRYENVKFNVDY
AVSGGFYGPPTAAELGGQFHHKSENGSVGAVFGAQVKK"
1852. .1855
1863. .4559
/gene="tbpA"
1863. .4559
/gene="tbpA"
/codon_start=1
/transl_table=11
/product="transferrin binding protein 1"
/protein_id="AAC43485.1"
/db_xref="GI:1055211"
/translation="MNKNLNLISLALSLFAVSQAQVQLNDVYVGTGKKAKHKE
NEYTGKVKVTPDSLSKEQVLGRDLTRYDPGIVSVVEQGRGATTGYSIRGVDRNVYG
LALDGLPOIQSVYVSOYSSRSSGAINIEYENLRSIQISKASSESGSGSLGSGVQFR
TKVSDIILKPGOSGWLDTKSAYSKNQWFLNSLAFAGTHNGFDALVLYTHRDSKETKA
KHDAESRSKIQRVLDLADNPQSNWFKVNDKPSLNCPEPQQVSYSDQKITYEQL
SAKEYTEERSLPDPVKYKSDSLWLVKLGYSKPKHYVAGVYEHSKQRYLDRDMTYPAY
WQLSDLKNGDKWTVMNNAKGLYRDNALDGVADIDYFTEDGVKSSKGLRWAKARIDEWH
TRDLGALYRYTNDQGNRLDRLSLDFDQOKINLSTRNENNGSEYPTIDKNGRATLD
KLWSSTKNEOSVEEKHDTIQLSLDKTVQGLKLNMLGLSDRNSLTKRHEILSE
FSVGTWHRIRGNGYKDPYIYELKDQAIYSKNEDYSCTIAGRADCATSKIKGNHVI
ALADNFATIKYLDIGLYRFDKHKFRSTRHWANQGDYKNSAWNIGIVAKTSELSLFI
RASGFRVPSFQELFLYDGMKSSDAYQKTEKLSPEKSLNQEAATFKGDFGVVE
VSYFKNYKOLIIAPERMHQTQSMINYNFVQDKLDGINLGLDNWGVDFKIPGEGY
TTLAYSKMRVKEQVNTQGMNIRSPILDTIQPARYVGVGYDQDEKWNGLMTWTHSS
GKNPNELRGNEQVFANERTATKRTSRWHTFDLTCYITPKHITTVRAGVYNLMNTR
YTTWESVROSSLNAIHQHTNVNDYAYAAAPGRNVVVSFEMKF"
BASE COUNT 1727 a 857 c 1074 g 1425 t
ORIGIN
Query Match 16.1%; Score 450.6; DB 2; Length 5083;
Best Local Similarity 52.4%; Pred. No. 1.9e-95;
Matches 1392; Conservative 0; Mismatches 1164; Indels 103; Gaps 14;
Qy 158 atgtctcgaacgtcctaaacgaagtaactggttggggaagtagtgaataatcagc 217
Db 1978 AAGCACATAAAGAGAACGACGACGCTTAGGGANAAGTAGTGAACACACGATT 2037
Qy 218 aaatgagtaaaaaatcaattcttgggtattcgatttaactcgtatgaccttgattt 277
Db 2038 CTCCTAGTAGGACGAGTGTAGGAATGCGAGATCTGACTCGCTACGATCCAGTATT 2097
Qy 278 cgggtggggaacaaggtcgcgctcaagtagtgccttcgcttcgattcgaggttaataaaa 337
Db 2098 CTGTAGTAGACCAAGGACGAGGTGCAACGACAGGCTACTCAATCTCGTGGGTATCGTA 2157
Qy 338 accgtgtcagcttacttgggtggtctaccacaagcgcaagttatcatcagctaggtt 397
Db 2158 ATCGTGTGGGCTTGGCATTAGACGGTTTGCACAGATTCATCTATGTAAGTCAATATT 2217
Qy 398 cagatgctaatggtggtgcaattaatgagattgagattgataaaacattcgttcaattgagt 457
Db 2218 CACGTTCTCTAAGCGGTGCCATTAATGAATAGATAACGAAAAATCTGGTTCGATCCAAA 2277
Qy 458 taagcaaaaggaagacttcgcggaataatgctctggtgcgcatggtggtgctattggtt 517
Db 2278 TTAGTAAGGAGCTAGTGTCTTCTGAGTTTGGCAGCGGCTCGCTAGGCGGCTCGGTGCAAT 2337
Qy 518 ttctactaaagatgctcagcagatattataaagggggcagcattggggcttagatagta 577
Db 2338 TCCGTACCAAGAGGTAAAGCAGCATTTAAGCCAGCGCAATCTTGGGAGCTAGATACCA 2397
Qy 578 agaccttattgcagcaaaaaatgccattttttacagttctatcgcgagcgctggtgagg 637
Db 2398 AAAGTGCTACAGCAGCAAAAAATCAACAATGGTTAAACTCACTTCTTTTCGGGTACTC 2457
Qy 638 cgggtggttttgaagcacttgatttcaactcaccacacacggttaagagagaccataatc 697
Db 2458 ACAATGGCTTTTGTGCTCTTGTGATTTACATCCCGTGTATGGTAAAGAAACGAAAGCTC 2517
Qy 698 attccgagggcaataataataataataattcggcgctataaccggccttgataaaatcgct 757


```
LOCUS      APL17915      5200 bp      DNA      BCT      02-MAR-2000
DEFINITION Actinobacillus pleuropneumoniae rhoAP gene and ORF 2 & 3.
ACCESSION  Y17915
VERSION    Y17915.1  GI:4127359
KEYWORDS   rho gene; rho protein.
SOURCE     Actinobacillus pleuropneumoniae.
ORGANISM   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
            Actinobacillus.
REFERENCE  1 (bases 1 to 5200)
AUTHORS   Tonpitak W., Thiede S., Oswald W., Baltes N. and Gerlach G.F.
TITLE     Actinobacillus pleuropneumoniae iron transport: a set of exbBD
            genes is transcriptionally linked to the tbpB gene and required for
            utilization of transferrin-bound iron
JOURNAL   Infect. Immun. 68 (3), 1164-1170 (2000)
MEDLINE   20143720
REFERENCE  2 (bases 1 to 5200)
AUTHORS   Thiede S.
TITLE     Direct Submission
JOURNAL   Submitted (11-AUG-1998) S. Thiede, Dept. of Microbiology and
            Infectious Diseases, School of Veterinary Medicine, Bischofsholer
            Damm 15, D-30173 Hannover, FRG
FEATURES   Location/Qualifiers
            1..5200
             /organism="Actinobacillus pleuropneumoniae"
             /strain="AP205"
             /serotype="7"
             /db_xref="taxon:715"
            271..1533
             /gene="rho"
            271..1533
             /gene="rho"
             /codon_start=1
             /transl_table=11
             /product="Rho protein"
             /protein_id="CAA76933.1"
             /db_xref="GI:4127360"
             /db_xref="SPTREMBL:O92AG7"
             /translation="MHLTQLNKTPVSDLVALGEGOMLENLARLRKODIIFAILKHOH
             KSGEDIFGOGILPDPGFLRSADSSVLGPDYVSPQIRRENLQTDGKIEGI
             RPKGEYERFALLKVLVDNDKPEVSRKILFENLPLHNSRLMERNGSTEDLTA
             RILDLAIPKGGQGLIVAPPKAGKTLLQNAQSITHHPECELLVLLIDRPEREVT
             EMQRTVGEVIASTFDEPATRHVQVAEMVIERAKRSVEHKDVVILLDSITRLRAYN
             TVTPSGKILSGVDANLHRPKRFCAARNVBEGSLTIIATALVDTGSKMDEVIFE
             EFKGTGMELHLSRKAERRVPAIEFRNSGTRKDDLMSPEHRNWLKVLNPM
             EVAAEWLIDKLGVAKTNEEFELMKRS"
             complement(1581..3518)
             /note="ORF2"
             /codon_start=1
             /transl_table=11
             /product="hypothetical protein"
             /protein_id="CAA76934.1"
             /db_xref="GI:4127361"
             /db_xref="SPTREMBL:O92AG6"
             /translation="MRKITPGDGIAPNPMKYQSGSWLFKLAYHFSPQHSVNAVFQ
             KQYDIDMAQARYSNLLPNNGKPEPRKKEVEYSLFGRNFYDDPDSOAQYGNLTNP
             SLVRSAYWRARFEDRHTKRLGFGYQVSKENDKGFEDRLNLNDQDIRLLTLHS
             LSCSRYPYDKNCKITPKRDANATYRNKYERKHQVOLSEKDLKIGKTOHLNLS
             MSYDRMSRHKRQDFERYLVYSKEDESEHPYIDVGNQDQVGYFNGSYKIPRIIGK
             RQIAVHNNTCKFDMHYGKKRDFSDCSAHTTGHQYAIGLSDRIFFGDKLDLGLVRY
             DHRFSDDDWNTGYTKNFSWNTGYTVRPREYIASSYRASNGFRVPSDFEMFGRVP
             GAADNYWRDLNKKLTSEKALNQEVGVAFNGNLGVLEVSYFDNRYRLIIVSAKEKG
             IDGIAGVNLQSLNLNGVNLKLNWHGIWSALPDGLYNLAIRNLKRVNVDTPKNVY
             FVRSSVLDALOPERYVASLGYDPNGKMGVNYLLTYSKAKKASELIGTYYTYGDIVK
             GYVNTQRFWYIDVGYTFAKHFTRVGRGIYLNLTNRKYSTWEAVRQTAETNVIPIYN
             RGHASVAPGRNVLSVAKF"
             complement(4385..4996)
             /note="ORF3"
             /codon_start=1
             /transl_table=11
             /product="hypothetical protein"
             /protein_id="CAA76935.1"
             /db_xref="GI:4127362"

BASE COUNT  1426 a 1191 c 948 g 1635 t
ORIGIN
Query Match      14.7%; Score 411; DB 2; Length 5200;
Best Local Similarity 51.3%; Pred. No. 3.7e-86;
Matches 1435; Conservative 0; Mismatches 1240; Indels 120; Gaps 16;

QY 72 atacgggtcgctgcgactgaaataaaaaatcgaagaaaaataacgatatagctgtcttcgga 131
DB 4329 ATCCGCTTTCCTGATTTTATCCGCACAAACTTATGCGAAGCAAAATGCAACGGAATTAGA 4270

QY 132 tgaagtattgtacagagagcgccattatgtctcagcaacgtcctcaaaaagaactgaactgctt 191
DB 4269 TAGCGTTTACAGTCAAAAGCAAAACGGAGATACACCGCCGACAAAACGAGTAACAGGTTT 4210

QY 192 ggggaaagttagtgaataattatcacgaaatgagtaaaaaatacaaatctcttggtattcgtga 251
DB 4209 GGGTAAAGTGACTAAATATCAGTAAAGATCTTGATAAAGACGAGGTACAAAATATTCTGTA 4150

QY 252 tttaactcgtatgacctgctgtatttcgggtggtggacaaggtcgcggtgcaagttagtgc 311
DB 4149 TCTGACCCGTTATGATCCCGGTATTTCGGTAGTTTGACCAAGGAAGAGCGCCAGTAGCGG 4090

QY 312 ctatgcattcggaggtgtagataaaaaacccgtgcagcttactgttctgtggtgggtaccaca 371
DB 4089 CTATTTCGATTGCGGGGTAGATCGCAATCGTGTGGCATTACTAGTAGACGGGTACAGCA 4030

QY 372 aggcacagttatcatcagctaggttcagatgctaatggtggtgcaataatgagattga 431
DB 4029 AATTCAGTCTCTATAAAGTCTTCTCTGTTTGACAAAGCGGAGCAATTAATGAAATCGA 3970

QY 432 gtatgaaacactcgttccattagtttaagcaaaagagcgaagttcgcggaatgagctc 491
DB 3969 ATATGAATAATTCGTTCGATTGAATTTAGTAAGGAGCGCGATTACAAAGAACGGTAG 3910

QY 492 tggtcgcacatggtgtgctattgtgtttctactaaagatgcgcagagatatattaaga 551
DB 3909 CGGTGCTTAGGCGGTGCCGTCCGCTCCGTACTATAAATCCCGACGATGTGATTAAAGA 3850

QY 552 ggggcagcattgggggttagatagtagaagaccttatgcagcaaaaaataagcattttt 611
DB 3849 AGGACAAAATTTGGGGATTAGACAGTAAACGGCGTATTTCGAGTAAATAAAGCAATTAG 3790

QY 612 acagttctatcgcagcgctgtgtagcggtgtgttttgaagcactgtttattgcaactca 671
DB 3789 CGAATCGCTTAGTTTTCAGGTGCTGTCGCCGGATTGGAAGCATTTGTTGCAATATAGTA 3730

QY 672 ccgacacggttaagagaccaaataattcattccgagggcaataataataataataatttcg 731
DB 3729 TCGAAAGGGCAAGAAAGCGGAAATTCATAAAGATGCGACT-AAGTAAGGCGAGTCCITCTA 3671

QY 732 gcgtataacgggtttgaaataatgcgtacgacttaccataatccgcacagatagtcctct 791
DB 3670 CCGTGTGTCGCCAAGAGAACTTAATTGGGATAAAAAATCAAAAGAGATAAATCGTTAGTCTT 3611

QY 792 ggaggatctcttttaattgtggaatacttcccaacattagattgtactcctcgtgc 851
DB 3610 TGAAGACCAATGCCACCACCGGGCAAGAGTAATCCCGGCTGTGTCCCGGCTCGGATTGT 3551

QY 852 aagggttaagttagaacgcgcgataatttcccagtgagaacatttcccggaatatatacgctga 911
DB 3550 ACAGGGGCGC-----AAAGTGATAAAAGATTTCGACTAATTCGGAAGA 3509

QY 912 agagcgcaaacagcttgagcagattctcttctcagctagcagctctcagcccaagaata 971
DB 3508 TTACACGAGGAGACGGCGGTATTGACCGCAATCCTATG----AAATATCAGTCGGGTTCTT 3452
```

QY 972 tacgggttaagatcgcatcgacacaccccttttagattacaagagtaattctgttttat 1031
Db 3451 GGTATTTAAATAGCTACCATTTCTACCGCAACATTCGGTAACGGCTATTGTAAC 3392
QY 1032 gaagtttggtatcaactcaactcgtctcattatcttggcgcaactcttgaagata--- 1087
Db 3391 AAACTAAGCAACGCTATGATATACAGGATATGCGACAAAGCAAAATATCAAAATTTATAC 3332
QY 1088 caaaacacgctcagatataccgtgatatgcaaacgcccagcttactatacaaaagacgata 1147
Db 3331 CGAATAATAAAGGAGAGCCGGATCGTAAAAAACCAGATGGAATATTCGCTATTCCGGTA 3272
QY 1148 ttaacttatcacttaggaactatgtttatgaagggaataataatttttagatggcttagtgt 1207
Db 3271 GAATTTCTATCCGGATGATCTTCTCAACCAATAGGT-----TATGGAATACGC 3221
QY 1208 tcaagccaagatcccttatgggtgctgctatagccatgtgaagtttttttgatgaacgctc 1267
Db 3220 TTAATCCTAGTTTAGTGCAGAGTGCATTATTGGAGCAGAGCGAGATTTTTTTGACGAACGGC 3161
QY 1268 accaacaacgctgttaggattcacctataataataaacacagaga---ataatcgctggt 1324
Db 3160 ATACGAACACCCCTTAGGGTTTGGTATCAGTATAGTCTGAAATGATTAAGAGCATTT 3101
QY 1325 tggatagcattaaactcagtcgagataaaacagatatgtaactatagccgctacatc 1384
Db 3100 TTGATCGTCTAAATCTCAATTTGGATCAACAGATATTCGATTACTGACCAAGTTACACA 3041
QY 1385 gcttcattttagcattatccctggtgtagataaaatccgcgcgacttttgataaaat 1444
Db 3040 GCCTAAGTTGTAGTAGTACCCTATTTTCGATAAAATTTGATAGATAACGCCGATAGC 2981
QY 1445 ctggtgtctatgtagaacgtagcgttaataataaccagaagaacat-----cg 1493
Db 2980 GAGATGCGAATGTACTTATGAACGGAATATTTATAGGAACGCTCATAAACAAAGTCCAAC 2921
QY 1494 tgcattcattagaatttgataaagcgctaaatgctggtcaaggcgctatttaacca--- 1550
Db 2920 TTTCTTTTGAAGAGATTGGAATCGGGAAACGCAACACTTATTATTAATTTATCCATGA 2861
QY 1551 -----aaccacaaactgaatttaggttggcttggatcgattta 1591
Db 2860 GTTATGATCGTATCGCTCTCAACCAATAGCGAGATCAATTTGAACGGCTTTATTAGTTT 2801
QY 1592 attcgctta tggatcatggtggtatgactgccccaat-----accaaaggcggttata 1645
Db 2800 ATTCTAAGGAAGATGAATCGGAACATCCGTATTATGATGTTGGAATCAAGATAAAGATG 2741
QY 1646 ccagctaccggttagagcggtttagataatccatatattatcg----- 1691
Db 2740 TCGGTTATTACAATAGCGGTACTTTATAATACCGGAGAAATATCGGTAAGAACAGCA 2681
QY 1692 -ccgcatccacgcagtatgaaacggtatcttctgtgtaataatacacgcggcgacatct 1750
Db 2680 TAGTCGCTCATATAATATTATGTAAGTTTGTATGCAATATGGAACGCAAGATTTTA 2621
QY 1751 taaactgtgaacgcgctaaattaaaggcgatagccattttgttagcttcgcgcgc--- 1807
Db 2620 GCGATTGTCGGACATATTACCAAAAGGCGAGCATTTATCGGATTTGGCTTATCGGATCGTA 2561
QY 1808 tagtgataacgagtagtggtgatttggttaggtggtgctttttgatcaacatcgattta 1867
Db 2560 TTTCTTCGTGTAAACTGGATCTAGGTTTGGGTGTACCGGTATGACTATCATCGTTT 2501
QY 1868 aattctgatgcgtggtgacacttagccgcaacttatcgaaatttggcttggaaatggtggga 1927
Db 2500 AATCGGATGATGTTGGTAAATACCGGTACATACAGAAATTTCTCTTGGAAACGGGAG 2441
QY 1928 ttaocgttaaaccaacagagtttgtatcgctttcttctatcgcattoaaacggttttagag 1987
Db 2440 TGACGGTTCGACCGAGAGAAATATATCGCTTCACTTACCAGGCTTCCAACGGTTTTAGAG 2381
QY 1988 tgcctgcattctatgaactttatggttaaacgctgatcatattgggtcttaagataaagaat 2047

Db 2380 TTCCGTCATTTGATGAATGTTCCGGTCGTCGAGTTC---CCGGTCCGCAGATAACAAT 2324
QY 2048 atgtgcaacgcgcgcacgtagccaccagtagagcagaaaaaatcgcataatcatyaga 2107
Db 2323 ATTGGCGTAGAGACTTACTTAAGAAAAAATAACTTCGGAAGAACACTTAACCAAGAAG 2264
QY 2108 ttggagtagctttaaaggtaatttggtttaaccttgatgtgagctatttccgtataaact 2167
Db 2263 TCGGAGTGGCGTTTAACGGCAATTTAGGAGTATTGGAAGTCAGTTATTTTGTATAACCGTT 2204
QY 2168 ataaaaatgatgtgagcagcatgtataaagaataatacaaaaaatcacactgtttctata 2227
Db 2203 ATCGACACTTGATGTGTCAGTGGCAAGAAAAAGGGATTGACGGAATAGCGGGTTTCTATA 2144
QY 2228 ataccataatattcaagatgtagcaactaaacgggataaatttagtcgtataattgact 2287
Db 2143 ATCTCCAATCGATTAAT-----CTTAATGGGGTAAACCTATTGGGTAAACTCAAT 2093
QY 2288 tacacggtattttatctatgctgcagatgggtttttttatcatcagttgcttataaacgtg 2347
Db 2092 GGCACGGTATTGGTGGCTTTACCGGACGGCTTATATTCAAATCTTGCTTATAACCGTA 2033
QY 2348 taaaaataaagagcggaactaaacgcgactcaagactcgatagcgtataacgcacttattc 2407
Db 2032 TTAAGTGAAGAACGTCGA---TACGAAGCCGAATTAACGTGTTGTACGTAGTCGGTAT 1976
QY 2408 tagatcgcttcaagccagcagcgtatgtctgttgatctcggtcgtacatccaccagaagaa 2467
Db 1975 TGGATCGGATACAGCCGGAACGCTATGTAGCCAGCTTAGGTTTATGATGAGCCGAACGTA 1916
QY 2468 aatggggaattggcattactaccacattctaaagccaaacacgcgcatgaggtggcag 2527
Db 1915 AGTGGGCTGTGAATATTATCTTAACCTACTTAAACGGAAGAAAGCCAGTGAATTAATCG 1856
QY 2528 gcaacgctcatcagcagcagcgtatgctgtttagtttagtgg-----caaacgacgggtt 2580
Db 1855 GTACGACGACTTACTACTTGGAGATGTGTTAAAGGTGTGACGAATCAACGTACTCGAC 1796
QY 2581 cttgttacacccatgatattaccgg---ttacatcaattataaaactcacaccttaccgtg 2637
Db 1795 CTTGGTATATTATGATGTTATTGGTTATTACACCTTTGCGAAACATTTTACCGTACGAG 1736
QY 2638 gaggaatttataatgactaataatcgttaataattccacttgggaatcagtcgccaatccg 2697
Db 1735 GTGGTATTATATACCAATCAACCAATCGTAAATACAGTACGTGGGAAGCGGTTCCGCCAACTG 1676
QY 2698 gtgtgaatgcagtaaaacgaacgcgggttagcaattcacactgatttggcgtcccgagg 2757
Db 1675 CCGAAACCAATGTGATCCGATTTAATTCGAGGTCTATCGGAGTTATGCCGACCGGGCA 1616
QY 2758 gaaatttcagtttagcatttgaagtgttag 2792
Db 1615 GAAATATGTACTAAGTGTAGAGCTAAATTTCTAG 1581

RESULT 8

U32780/c

LOCUS

DEFINITION

U32780 L42023

ACCESSION

U32780.1

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 10205)

Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A.,

Kirkness,E.F., Kerevage,A.R., Buit,C.J., Tomb,J., Dougherty,B.A.,

Merrick,J.M., McKenney,K., Sutton,G.G., FitzHugh,W., Fields,C.A.,

Gocayne,J.D., Scott,J.D., Shirley,R., Liu,L.I., Glodok,A.,

Kelley,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T., Hedblom,E.,

BCT

29-MAY-1998

DNA

10205 bp

Haemophilus influenzae Rd

Haemophilus influenzae Rd

Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

Haemophilus.

1 (bases 1 to 10205)

Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A.,

Kirkness,E.F., Kerevage,A.R., Buit,C.J., Tomb,J., Dougherty,B.A.,

Merrick,J.M., McKenney,K., Sutton,G.G., FitzHugh,W., Fields,C.A.,

Gocayne,J.D., Scott,J.D., Shirley,R., Liu,L.I., Glodok,A.,

Kelley,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T., Hedblom,E.,

Cotton, M.D., Utterback, T., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Geoghagen, N.S., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O. and Venter, J.C.
Whole-genome random sequencing and assembly of Haemophilus influenzae Rd

Science 269 (5223), 496-512 (1995)
93350630

JOURNAL
MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS


```

NGETGLISQFSADFGTKKLKGLFYTKRKTNNQNYEKKKLYDIDANIYSNRFGRKVK
PTKDEEHPFTREGTLEGGFYGPNGEELGKFLAGDKKVFVGSFAKEEYKOKTLP
KETLIDGKLTFTSTKPDATSTTNAKTDATNAENFTTKDISSEGEADYLLIDNP
VPLLPTEUNGDPATSKHVEVRDKTYKVEACKNLSYVFGMYEDENKKNKNEREQY
HOLLGLRTASSKIPGTGNVYKWSFGYISDGETSYTGDGRDKDNVAFEDVNF
EKLKGLLRKADQNPVFIEANFNKNGNAFTGTATKDLVLDGKNSQTKNPTINITT
KVNGAFYGNASELGLGYFTYNGKNPTDKNSPTASSPSNSEKARAAVVFAGKQVETNN
K"
8642..9082
/gene="HI0996"
CDS
8642..9082
/feature="H. influenzae predicted coding region HI0996"
/feature_id="AAC22658.1"
/db_xref="GI:1574027"
/translaton="MNGVKOLFLLSLGLSLNVAAEVARPKNDTLNTIOSAELK
TSFSFSPKKEIPNRHIIISLSKSLAHPLRLVRLGLIPALYQNNQVALLPLYKQF
PQDNFLTLWAKAIEAREODLQTSIAYIRELFARDASLLPLRY"
9398..10087
/gene="HI0997"
CDS
9398..10087
/feature="H. influenzae predicted coding region HI0997"
/feature_id="AAC22659.1"
/db_xref="GI:1574026"
/translaton="MNGNGKYVDNKKYNENATVRIGGGLGYGTAVESVLFPFQEKR
WYAGSSGNTMKQYQKDLGIRLENVDLSKTDQIISALEYGESRYKIRKHLGDNIFY
VSTLFLPKSTQFVFGMDFHRENTQALDNYOQKTLRLGWQDFYFYGISRLTFSY
ANRYREKDLIGIQKNREYTTITLHNRNHPMGLTPKLSWDYQKSTNSHAFYRDK
NRYLIGIKIF"
BASE COUNT      3008 a 1866 g 1840 g 3491 t

Query Match      12.4%; Score 345.2; DB 2; Length 10205;
Best Local Similarity 51.2%; Pred. No. 1.1e-70;
Matches 1370; Conservative 0; Mismatches 1123; Indels 181; Gaps 17;

Qy 173 aaacgaagtaactggctggggaagtagtgaataattatcacgaataatgataaaatc 232
Db 6449 ATAATGAAGTAACATGGACTTGGCAACATTAATCAAAACTAGTGAAGATATCAGCCGGAAC 6390

Qy 233 aaattcttggtattcgtgatttaactcgctatgacctggtatttcggtggtggaacaag 292
Db 6389 AAGTATTAAATATTCGTGATCTAACACGCTATGATCAGGGATTTTCAGTTGTACAACAAG 6330

Qy 293 gtcgcggtgcaagtagtgctatgccattcgaggtgtagatataaaacccgtgctagcttac 352
Db 6329 GTCGCGGTGCAAGTCTCGGATATTCATTCGTGTATGGACAGAAATAGATTGCTTTAT 6270

Qy 353 ttggtgtaggctaccacaagcgacaggtatcatacagctagggttc----- 398
Db 6269 TGGTAGATGTTTACCTCAACAGCAATCTTATGTAGTGAAGCCCTTTAGTTGCTCGTT 6210

Qy 399 ----agatgctcaatgggtggtcaatgaattgagattgagtagaacaattcgttcaattg 454
Db 6209 CAGGATATTCGGCACATGTCATTAATGAATTTGAATATGAATATGAATATGAAGCCGCTCG 6150

Qy 455 agttaagcaaaaggagcaagttctcggaatatggctcgtggtgcgcattggtggtgctattg 514
Db 6149 AAATAAGCAAGGGGGAGTCTTCTGAGTATGGAAATGGAGCATTAGCTGGTCTCTGTAA 6090

Qy 515 gtttctgactaaagatcgccaggtatattataaaggaggcgacgattggggcttagata 574
Db 6089 CATTTCAAGCAAAATCCGCACGCCATCTCTTGAAGGAGAGCAAAATCATGGGGAAATTCAAA 6030

Qy 575 gtaagacctcttattgccagcaaaaatagccatttttttacagtctatcgacggcggtg 634
```

```

Db 6029 CTAAAAATGCTTATTCAAGCAAAAAATAAGGCTTTACCCATTCTTTAGCTGTAGCAGGAA 5970
Qy 635 aggcgggtggttttgaagcaactgttattgcaactcacgacacggttaagagagaccacaaa 694
Db 5969 AACAAAGTGGATTTGAAGGACTTTGCTATTATTACACTCAACGAAATTCATTTGAACCAAG 5910
Qy 695 ttcattccagagcaataataataaataataattccgggtataaacccggctttgaaatc 754
Db 5909 TCCATAAAGATGCTATAAAGGGCTGCAAAAGTTATTAATGCTTTAATCGCC----- 5858
Qy 755 gctacgactttaccacaaatccgcacagaaatgctcctggaggatcctccttttaattgtg 814
Db 5859 -----AAACCAGAGAAATCAATTCGCATACTTTG-----TGATGG 5826
Qy 815 aagatcttggcccaacattagattgtactcctcgtgcaaggggttaagttgaaacgagata 874
Db 5825 AAGATGAGTGTCCA-----AAGGGTTATGACGAGTGTATACC 5789
Qy 875 atttccagtgagaacattttccgggaataatacgcctgaagagcgcaaacagcttgagcaga 934
Db 5788 TTCAGC-----CAACCACTTCGCGATT 5766
Qy 935 ttccttctcagcagcagctcctcagcccaagaataataccggttaagatcgcattgcac 994
Db 5765 TATCCACCAAAAAAGAACCGTAAGCGTTTCAGATTATACGGGGCTAACCGTATCAAAC 5706
Qy 995 caaacctttagattacaagagtaattctgtttttttagaagtttggctatcaacttcaact 1054
Db 5705 CTAATCCAATGAATATGAAGCCAGTCTTTGGTTTTTAAGAGGAGGTATCATTTTTCTG 5646
Qy 1055 cgtctcatctctggcgcaactctagaagatacaaaaacacgctacgatatccgagata 1114
Db 5645 AACAACTATATATGGTGGTATTTTGAATTCACACACAAAAATTTGATATCCGCGATA 5586
Qy 1115 tgcaaacgcagcttactatacaaaaagacgatatctaacttatcacttaggaactatggtt 1174
Db 5585 TGACATTTCCCGCTTATTTAAGCCCAACAGAAAAAGGGGATTTAGCAAAATCGCCCTTTT 5526
Qy 1175 atgaaggggataatttttagatggcttagtggttcaagccaagatccctta1gggtgc 1234
Db 5525 ATCCAAAGCAAGAT-----TATGGCGCATATCAACATATATTAGGATGGCGAGCGGCTTA 5472
Qy 1235 gctatagccctggaagttttttagaagcgtccacacaacgctcgttttaggattcacct 1294
Db 5471 AATATGCAAGTGGCTTATTTTCGATGAACACCATAGAAACACGCGTGTAGTATTGAAT 5412
Qy 1295 ataaataaaaccagagaataa---tcgctggttggtatagcattaaactcagtgcgagata 1351
Db 5411 ATATTTACGAAAAATAAGAACAAAGCGGGCATCATTTGACAAAGCAGTGTAAAGTCTAATC 5352
Qy 1352 acaagatatgtaactatagccggtacatcgcttgcattgttagcgattatctctgtg 1411
Db 5351 AACAAACATCATATCTTGACAGTTATATGCGACATACGCAATTCGAGTCTTTTCTCTAATC 5292
Qy 1412 tagataaaattcccgcaactttggataaacttgggtctatgtatcgaactgagcgata 1471
Db 5291 CAAATAAGAAATTCGCGCCCAACACTTGTATAAACCCTTATTCTACTATCTGTTCTGATAGAA 5232
Qy 1472 ataattaccaagaaaagcagctgctcattcatttagaatttgataaa-----gcgctaa 1525
Db 5231 ATGTTTATAAAGAAAAACATAAATATGTTGCAATTTGAATTTAGAGAAAAAATAATCAACAAA 5172
Qy 1526 atgctggtgcaagcgctatttaacaaacccacaaactgaatttaggggttgggcttgcac 1585
Db 5171 ATTGGCTTACTCATCAAAATTTGCTTCAATCTTGGTTTTGATGACTTTTACTTCCGCACTTC 5112
Qy 1586 gatttaattcgtctatgatacatgagggtatgactgccccataataccaaagggcggtcata 1645
Db 5111 AGCATAAAGAAATTTAACTCGACGTTTATTCGCTACGGCAGATAGTATTTTCAGATATAAA 5052
Qy 1646 ccagctaccggtgtagaggcggttta--gataatccatataatttattcgccgagctccacgc 1704
```


[illegible]

BASE COUNT	1834 a	849 c	960 g	1456 t
ORIGIN				

Query Match	12.2%;	Score 340.8;	DB 2;	Length 5099;
Best Local Similarity	50.4%;	Pred. NO. 1.2e-69;		
Matches 1347;	Conservative	0;	Mismatches 1152;	Indels 175;

3 14;

Qy	173	aaacgaagtaactcgtgcttggggaagtagtgaaaaataatcacyaaatagtgataaaatc	232
Db	2342	ATATGAAGTAATCTGACATTGGCAAAATATCAAAACCTAGTGAAGATATCAGCCGAGAAC	2401
Qy	233	aaattcttggtattcgtgatttaactcgtatgcctcgtgatttcctggtggtgaaacaag	292
Db	2402	AAGTATTAATAATTCGTGATCTAACACGCTATGATCCAGGCATTTAGTTGTAGAACAAAG	2461
Qy	293	gtcgcggtgcaagtagtggtcgtcattcggagtgtagataaaaaacggtgcagcttac	352
Db	2462	GCCGTGTCGAAGTTCTGGATATCTATTCTGTGGTATGACACAGAAATAGAGTTGCTTTAT	2521
Qy	353	ttgttaatgggtcacacaagcgacagtatacaacgctagggtc-----	398
Db	2522	TAGTAGATGGTTTACCTCAAGCGCAATCTTTATGTAGTCAAAAGCCCTTTAGTTGCTCGTT	2581
Qy	399	----agatcctaagtgtgtgcaataatagattgagattgagataaaaaccttcgattcaattg	454
Db	2582	CAGGATATCTGGCACTGGTGCAATTAATGAATTAATGAATTAATGAATTAATGAATTAATGA	2641
Qy	455	agttaagcaaaagcagaagttcttcgggaatatggctctggtgcgcagtgtggtgctattg	514
Db	2642	AAATAAGCAAGGGGGGAGTCTTCTGAGTAGTGTAATGGAGCACTAGCTGGTTCTGTAA	2701
Qy	515	gttttcgtactaaagatgcgcggatattattaaagagggcagcatttggggccttagata	574
Db	2702	CATTTCAAAAGCAATCAGCAGCCGATATCTTAGAAGAGAGACAAATCATGCGGAATTCAAA	2761
Qy	575	gtaagaccttatgcacgcaaaatagccattttttacgtctatcgcagcgcgctggtg	634
Db	2762	CTAAATGCTTATTCACGCAAAATAAAGGCTTTTACCCTATCTTTAGCTGTAGCTGGAA	2821
Qy	635	aggcgggtggttttgaagcacttgttattgcaactcacccgacacggttaaagagaccaaaa	694
Db	2822	AACAAGGGGATTTGACGGGTCGCCATTTATCTACACAGAAATCAATTGAACCCCAAG	2881
Qy	695	ttcattccagaggcaataataataacaataatattcgcgcgtataaacoggctttgaaaaatc	754
Db	2882	TCCATAAAGATGATTTAAAGGGGTACAAAGTTATCATCTGATTAATGCC-----AAACC	2936
Qy	755	gtcacgactttaccaaaattccgcacagaatgctcctcgaggatctccttttaattgtgg	814
Db	2937	AGAGGATCAATCTGCATACTTTGTGATGCAAGATGAGTGTCCAAGCCAGATGATTAATAA	2996
Qy	815	aagatacttgcaccaacattagattgtactcctctgcgaaggggttaagttgaaccgcgata	874
Db	2997	CAGTGTGTTTACCTTTC-----	3012
Qy	875	atttcccaagtgaacaatttccgggaatacaocctcgaagcgcgcaaacacgcttgagcaga	934
Db	3013	-----GCCAAACGACCTCGCATTT	3031

Qy	935	ttccttatcgac	ctgagcgtctctgaagcccaagaataataccgggtaagatcgcatttcgac	994
Db	3032	TATCCTCCCAAGAGAAAACCGTAAGCGTTTCAGATTATACGGGGCTAACCGTATCAAAAC		3091
Qy	995	caaaccttttagattacaagagtaattcgttttttatgaagtttggctatcaactcaact	1054	
Db	3092	CTAATCCAAATGAATAATGAAGACCGAGCTGTGTGTTTTAAGAGAGGGTATCATTTTTCGT	3151	
Qy	1055	cgctcattatctctggcgcaactcttagaagatacaaaaacacgcctacgatatccgtgata	1114	
Db	3152	AACAACATATATGTGTGTTATTTTGATTCACACACAAAAATTTGATATCGTGATA	3211	
Qy	1115	tgaaagccagcttactatacaaaagacgataattaacttatcacttaggaactatgttt	1174	
Db	3212	TGACATTTCCCGCTATTTAAGATCAACAGAAAAACGGGATGATAGCAGTGGCTCTTTTT	3271	
Qy	1175	atgaagggtataatttttagatggccttagtttcaagccaaagatcccttatgggtgc	1234	
Db	3272	ATCCAAAGCAAGATATATGTGCAATATCAACGATTTAGAGATGGCG-----AGCGGTTA	3325	
Qy	1235	gctatagccatgtgaagtttttttgatgaacgtcaccaaaacgcgtctttaggattccact	1294	
Db	3326	ACTATGCAAGTGGCTTTATTTTCGATGACACCATAGAAACACGGCTGTAGGTATGAT	3385	
Qy	1295	ataaatataaccagagaataatcgctggttggtagcattaaa---ctcagtcgagata	1351	
Db	3386	ATATTTACGAAAAATAAGCAAAAGCGGCATCATTTGACAAAGCAGTGTAAAGTGCTAATC	3445	
Qy	1352	aacaagatatgaactatatagcggcgtacatcgcgttcgattgtagcattatcctgtgg	1411	
Db	3446	AACAAAAACATCATCTGTGACAGTTATATGCAACATACCGATTTGCAGTCTTTATCCTAATC	3505	
Qy	1412	tagataaaatttcgcccgcgacttttgataaaatcttggctctatgctcgcaactgagcgta	1471	
Db	3506	CAAGTAGAATTCGCCGCCACACACGCTGATAAACCTTATTCATACTATCATTTCTGATGAA	3565	
Qy	1472	ataattaccaagaaagcatcgtgctatcattagaatttgataaagcgtataatgctg	1531	
Db	3566	ATGTTTTATAAGAAAAACATAATATGTTGCAATTTGAATTTAGAGAAAAAAATTCACAAA	3625	
Qy	1532	gtcaagcgctatttaaccaaaaccacaaactgaatttaggg-----ttggggcttgg	1582	
Db	3626	ATTGGCTTACTCATCAAAATGTGCTTCAATCTTGGTTTTGATGACPTTACTTCAGCGGCTC	3685	
Qy	1583	atcgatttaatcgcctatggatcagtggggatagactgcaccaatataccaaagcggtt	1642	
Db	3686	AGCATAAAGATTAATTAACTCGACGTGTTCACCGCTACGCCAAGAGTATTTACAGAGAAAG	3745	
Qy	1643	ataccagctaccgcggtagaggcgcttagataatccatattattatcgccgcgcatccac	1702	
Db	3746	CTAATGAACAAGAAGAAATGGTTACAAAAACAACCTTACTTATACCCAAAAACCAACAG	3805	
Qy	1703	gcagtatgaaacggttatcttctgttaataatacacg-----cggcgaactcttaaac	1756	
Db	3806	TAGTTTTGTAGTACAAGATCATTTGTGATTAATAAGGTAACCTCTTAATTACAGAGACT	3865	
Qy	1757	gtgaacgcgtataaataaagcgtatagccatttctttagcttcgcg---gactagtga	1813	
Db	3866	GTAAGTTCGGTTAATTAAGGGAAAAATTAATNTTTCGACGACGCAATTAATATGGCAT	3925	
Qy	1814	taagcgagtatgtgattttgggattaggggctgttttttgatacaacatcgatttaaatctg	1873	
Db	3926	TAGGAAATACGTTGATTAGGTTTAGGTATTCGGTATGACGTATCTCGCACAAGACTA	3985	
Qy	1874	atgatacgtggacacttagccgaacttatcgaaattggtcttggaaatggtgggattacgc	1933	
Db	3986	ATGAATCAACTATTAGTGTGTGTTAAATTTAAAAATTTCTCTGGAAATACTGGTATGTCA	4045	
Qy	1934	ttaaaccaacagagatttgatcgctctcttctacgcatctcaaacggttttagagtgc	1993	
Db	4046	TAAACCCACAGGAATGGCTTGATCTTTCTTATCGCCTTTTTCTACTGGATTTAGAAATCCTA	4105	
Qy	1994	cattctatgaactttatggttaaacgtgatactattggccttaagataacgaatatgtgc	2053	

```
Db 4106 GTTTTGGCTGAATGATGTTGGCGGTATGGTGGCAATAATA----- 4147
Qy 2054 aacgcgcgaacgtagccaccagttgagccagagaaaaatcgactaatcatgagattggag 2113
Db 4148 ---GCGAGGTTTATGTAGTAAATTAAGCCTGAACACATCTCGTAACCAAGATTGTGTC 4204
Qy 2114 ttgattttaaaggtcaattggttacccttgatgtagctatttcogtataaactataaa 2173
Db 4205 TCGCTCTAAAAGGGGATTGTGGTAATATGAGATCAGTCATTTAGTAAATGCTATCGAA 4264
Qy 2174 atatgtgacgacgacatgtaaa---agaataatacaaaatcacactgtttctataact 2230
Db 4265 ATCTTATGCGCTTTGCTGAGAACCTTATAAATAATGGAACCTGGAAAGGCCAATTATGAT 4324
Qy 2231 accataattattoagatgtagcactaaacgggataaaatttagtcgctaaatttgacttac 2290
Db 4325 ATCATAATGCACAAAATGCAAAATAGTTGGGTAAATATAACTGCGCAATTAGATTITA 4384
Qy 2291 acggtattttatctatgctgcagatggtttttatttcattcatcagttgcttataacogtcaa 2350
Db 4385 ATGGTTTATGGAACAGTATTCCTACGTTGGTATGCAACATTTTGTCTATAACCGAGTAA 4444
Qy 2351 aagtaaaagcggaacactaacggaactcaagactcgaagctgatagcgtaaacgatcctattctag 2410
Db 4445 AAGTTAAAGATC---AAAAATCAATGCTGGTTGGCCTCCGTAAAGCAGTTATTATTATTG 4501
Qy 2411 atgcgattcagcgcacgctatgctgtgattcggctagcagatcacccagaagaataa 2470
Db 4502 ATGCCATTCAGCCACGCGTTATATCATTTGGTTAGGTATGATCATCCCAAGTAATACATT 4561
Qy 2471 ggggaattgycattactaccacctattttaaagccaaacacgcogagtgagtgccagca 2530
Db 4562 GGGGAATTAATACAAATGTTTACTCAATCAAAAGCAAAATCTCAAAATGAATTGCTAGGAA 4621
Qy 2531 caggtcatcagcgatacatcgc-----gttgattagggtggcaaacactgaccggttctt 2583
Db 4622 ACGTGCATTTGGGTAAACAATTCAGGGATGTAAATCAACAAGAAATCTTACTCGGGCAT 4681
Qy 2584 ggtacaccatgatattaccgg---ttacatcaattataaaactacacctttagctggag 2640
Db 4682 GGCATATCTTAGATGATTCGGGTATTACATGGCGCAATTAATAATATTATGCTTCGATTAG 4741
Qy 2641 gaattataatgactaatcgttaaatattccacttgggaatcagtcagtcgccaatccggtg 2700
Db 4742 GGATATAAATTTATTCAACTATCCTATGTTACTTGTGGAAAGCGGTGGCTCAAAACAGCAC 4801
Qy 2701 tgaatgcagtaaaccaaacgacgg---gtagcaattacactcagattggcgctccgggga 2757
Db 4802 AAGGTGCGGTCAATCAACATCAAAATGTTGGTAGCTATCTCGTACGACGATCAGGAC 4861
Qy 2758 gaaatttcagtttagcatttgaatgaagtttta 2791
Db 4862 GAAACTATACCTTAACATTAGAAATGAATCTA 4895

RESULT 10
LOCUS I79990 5099 bp DNA PAT 10-JUN-1998
DEFINITION Sequence 4 from patent US 5708149.
ACCESSION I79990
VERSION I79990.1 GI:3208280
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 5099)
AUTHORS Loomore,S., Harkness,R., Schryvers,A., Chong,P., Gray-Owen,S., Yang,Y., Mordin,A. and Klein,M.
TITLE Method for producing purified recombinant Haemophilus influenzae transerrin binding proteins
JOURNAL Patent: US 5708149-A 4 13-JAN-1998;
FEATURES Location/Qualifiers
```

```
source 1. 5099
BASE COUNT 1834 a 849 c 960 g 1456 t
ORIGIN
Query Match 12.2%; Score 340.8; DB 5; Length 5099;
Best Local Similarity 50.4%; Pred. No. 1.2e-69;
Matches 1347; Conservative 0; Mismatches 1152; Indels 175; Gaps 14;
Qy 173 aaacgaagtagcgttgggaaagtagtaaaattatcacgaatgtaaaatc 232
Db 2342 ATAATGAAGTAACCTGGACTTGGCAAAATATCAAAACTAGTGAAGTATACGCGAGAA 2401
Qy 233 aaattcttggctatcgatttaactcgctatgacctggtatttcggtggtgaacaag 292
Db 2402 AAGTATTAAATATTCGTGATCTACACGCTATGATCCAGGCATTTTCAGTTGTGAACAA 2461
Qy 293 gtccggtgcgaagtagtggctatgccattcaggtgtagataaaacccgtgtcagcttac 352
Db 2462 CCCGTGGTGAAGTTCGTGATATTCTCGTATGTTGATGGACAGAAATAGAGTTGCTTTAT 2521
Qy 353 ttgttgatggctaccacaagcgacagttatacagctaggttc----- 398
Db 2522 TAGTAGATGGTTTACCTCAAAACGCAATCTTATGTAGTGCAAGCCCTTTAGTTCGCTT 2581
Qy 399 ---agatgctcaatggtggtgcaatttaagattgagttgagtaaaacattcgttcaattg 454
Db 2582 CAGGATATTCGGGCACATGGTGCATTTAATGAATTTGAATATGAATAATGAAGGCCGTCG 2641
Qy 455 agttaacgaaggagcgaagttctcgcgaatatgctcctggtgcgcattggtggtctattg 514
Db 2642 AAATAAGCAAGGGGGGAGTTCTTCTGAGTATGTAATGGAGCAGTACTAGTGGTCTGTAA 2701
Qy 515 gtttcgtactaaagatgcgcaggtatatttaaagggggcagcattggggcttagata 574
Db 2702 CATTTCAAGCAAAATCAGCAGCCGATATCTTTAGAAAGAGAGACAAATCATGCGGAATCAA 2761
Qy 575 gtaagacctcttagccagcaaaaatagccattttttacagctctatcgacgcgctgtg 634
Db 2762 CTAAATATGCTTATTCAGCAAAATAAAGGCTTTACCCATTCCTTACCTGTAGCTGGAA 2821
Qy 635 agggcggtggttttgaagcactgttattgcaactcaccgcacgcggttaaaagagaccaaaa 694
Db 2822 AACAAAGGGGATTTCAGCGGTGCGCATTTTATCTCAACGAAATTCATTTCAAAACCCCAAG 2881
Qy 695 ttcaatccgaggaacaaataaataaaacataatttcggcgataaacggcgtttgaaatc 754
Db 2882 TCCATAAAGATGCAATTAAGGGCGGTACAAAGTTTATCATCGATTAAATCGCC-----AAACC 2936
Qy 755 gctacgactttacccaaattccgcacagaatgctccctggagatcctcttttaattggtg 814
Db 2937 AGAGGATCAATCTGCATACCTTTGTGTGATGCAAGATGATGTCCAAAGCCAGATGATTATA 2996
Qy 815 aagatacttgcaccaacattagattgtactcctcgtgcaagggttaagtgaacgcgcgata 874
Db 2997 CAGTTGTTTACCTTTC----- 3012
Qy 875 atttccacgtgagaaacatttccgcgaatatatacgcctgaagagcgcaaacagcttgagcaga 934
Db 3013 -----GCCAAACGACCTGCGATTT 3031
Qy 935 ttccctatcgactgagcagctctcagcccaagaatatacgggttaagatcgctattgcac 994
Db 3032 TATCTCCCAAGAGAAACCGTAGCGCTTCAGATTATACGGGGGCTTAACCGTATCAAAAC 3091
Qy 995 caaaccttttagattacaagaagtattctgttttttgaagtttggcttaccactcaact 1054
Db 3092 CTAAATCAATGAATATATAAGCCAGCTTGGTTTTTAAAGAGGAGGTTATCATTTTCTG 3151
Qy 1055 cgtctcattatcttggcgcgaatctttagaagatacaaaaacacgcgtacgatccgtgata 1114
Db 3152 AACAAACATTATATTGGTGGTATTTTGAATTCACACAAACAAATTTTGATATCGTGATA 3211
```

Qy 1115 tgcaagcgcagctactatcaaaagacgatattataacttacttaggaactatgttt 1174
Db 3212 TGCAATTTCCCGCTTATTTAAGATCAACAGAAAAACGGGATGATAGCAGTGGCTCTTTT 3271
Qy 1175 atgaaggaggaataatttttagatggcttagtggttcaagcccaaggatcccttatgggtgc 1234
Db 3272 ATCCAAACAGATATATGGTCATATCAACGTATTGAGGATGGCCG-----AGCGCTTA 3325
Qy 1235 gctatagccaatgtgaagttttttgatgaacgtcaccaaaacgcgtcttagattaccct 1294
Db 3326 ACTATGCAAGTGGGCTTATTTCGATGAACACCATAGAAAAACAGCGTGTAGGTATTGAAT 3385
Qy 1295 ataaataataaccagagaataatcgcgtgtgttgatagcattaaa---ctcagtcgagata 1351
Db 3386 ATATTTTACGAAAAAAGAAACAAAGCGGCATCATTTGACAAAAGCAGTGTAAAGTCTTAATC 3445
Qy 1352 aacaagatatgaactatatacgccggtcacatcgtctgtcattgttagcgattatcctgtgg 1411
Db 3446 AACAAACATCATACTTGCACAGTTATATGCAACATACGCATTGCGACTTTTATCCTTAATC 3505
Qy 1412 tagataaaattgcgcgcgactttggataaattcttggctctatgtatgcgaactgagcgta 1471
Db 3506 CAAGTAAGAATTGCGGCCAACACGCTGATAAACCTTATTCTACTATCATCTCTGTAGAA 3565
Qy 1472 ataattaccagaagaacatcgtgtcattcattttagaatttgataaagcgcataaactcgt 1531
Db 3566 ATGTTTATAAGAAAAACATAATATGTTGCAATTGAATTTATAGAAAAAATAATCAACAAA 3625
Qy 1532 gtcaaggcgtatttaacaaaccccaaacactgaatttaggg-----ttggccttg 1582
Db 3626 ATTGCTTTACTCATCAAAATGCTTCATCTTGTTGTTTGATGACTTTACTTCACGGCTTC 3685
Qy 1583 atcgatttaattcgtcttattgatatcgtgggatactgactgcccaataataccaaaagcggtt 1642
Db 3686 AGCATAAAGATTATTAACTCGCGTGTACCGCTACGCAAGAGTATTTCAGAGAAAG 3745
Qy 1643 ataccagctacgcggttaggggcttagataatccataatttatccgcgcgcatccac 1702
Db 3746 CTAAATGAACAAAGAAAGAAATGTTACAAAAAACAACTTACTTATACCAAAACCAACAG 3805
Qy 1703 gcagttatgaacggtatcttctgtgaataatacag-----cggcgacatcttaact 1756
Db 3806 TAGGTTTTGTAGTACAAAGATCATCTGTGATATTAAGGTAACTCTCTTAATTACAGAGACT 3865
Qy 1757 gtgaaccgcgtaaattaaagcgatagccattttttagcttcgc---gatctagtga 1813
Db 3866 GTAAAGTGGGTTAATTAAAGGGAATAATTATTATTCGCAGCAGCAATAATAATGGCAT 3925
Qy 1814 taagcgagatagtggatttggattaggggtgcgttttgatcaacatcgatttaaatcgt 1873
Db 3926 TAGGGAAATACGTTGATTAGGTTTAGGTATTCGGTATGACGTATCTCGCACAAAAGTA 3985
Qy 1874 atgacctgtgacacattagccgaacttatcgaaattgctcttggaaatggtgggattacgc 1933
Db 3986 ATGAATCAACTATTAGTGTGTAAATTTAAAAATTTCTCTTGGAACTACTGTTATGTCA 4045
Qy 1934 ttaaacacacagagtgttatcgctcttcttatcgacttcaaaaggttttttagagtgctt 1993
Db 4046 TAAACCAACGAAGTGGCTGTATCTTCTTATCGCTTCTTCTACTGGATTAGAAATCCTA 4105
Qy 1994 cattctatgaactttatggtaaacgtagatcatattgggcttaagataaacgaatatgtgc 2053
Db 4106 GTTTTGTCTGAAATGATGTGTTGGCGGTATGTTGGCAATAATA----- 4147
Qy 2054 aacgcgcgaacgtagccacacgtaggcgagaaataatcgactaatcatgagattggag 2113
Db 4148 ---GGCAGGTTTATGTAGTAAATTTAGCCCTGAACACATCTCGTAACCAAGAGTTTGGTC 4204
Qy 2114 ttgattttaaagggtcaatttggttacccttgatgtgagctatttccgttaataactataaaa 2173
Db 4205 TCGCTCTAAAGGGGATTTGGTAATATTGAGATCAGTCACTTTTAGTAAATGCTTATCGAA 4264

Qy 2174 atatgattgcagcagcatgtaaa---agaataatacaaaaaacacactgtttctataact 2230
Db 4265 ATCTATTCGGCTTTGCTGAAGAACTTAATAAATAATGAAGTGAAGGCCAATATATGGAT 4324
Qy 2231 accataatttcaagattagcactaaacgggataaaatttagtcgcgtataaatttagcttac 2290
Db 4325 ATCATATGCAAAAATGCAAAATTAGTTGGCGTAAATAATAACTGCGCAATTAGATTTTA 4384
Qy 2291 acggtatttatctatctgctccagatgggttttttattcatcagttgtctataaacggttaa 2350
Db 4385 ATGGTTTATGAAACGTTATTCCTACGTTGGTATGCAACATTTGCTTATAACCGAGTAA 4444
Qy 2351 aagtaaaagcgggaaactaacgcgactcaagactcgatcgtaaacgactcctattctag 2410
Db 4445 AAGTTAAAGATC---AAAAATCAATCGTGTGTGGCTCGTAAAGCAGTTATTATTG 4501
Qy 2411 atgcgattcagccagcgcgtatgtcttgattcgtctacgaccccccgagaagaaaaa 2470
Db 4502 ATGCCATTACGCCAGCCGTTATATCATTTGTTTAGCTATGATCATCCAAGTAATACTT 4561
Qy 2471 ggggaattgtgcattactaccacattcttaagccaaaacgcgcgatgggtgagcgca 2530
Db 4562 GGGGAATTAATACAAATGTTTACTCAATCAAAAGCAAAATCTCAAAATGAATTGTAGGAA 4621
Qy 2531 cagtcacacgagatacatcgc-----gttgatttaggtggcaaacactgacggttctt 2583
Db 4622 AACGTGATTTGGGTAAACAATTCAGGGATGTAAATCAACAAGAAACCTTACTCGGGCAT 4681
Qy 2584 ggtacacccctgatcattccgg---ttacatcaattataaaaaactacaccttacctgtgg 2640
Db 4682 GGCATATCTTAGATGTATCGGTTATTACATGGCGAATAAAAATATTATGCTTCGATTAG 4741
Qy 2641 gaattataatgtgactaatcgttaaatattccacttgggaatcagtgcgcaatcccggtg 2700
Db 4742 GGATATATAATTTATTCAACTATCGCTATGTTACTTGGGAAGCGTGCCTCAAAACAGCAC 4801
Qy 2701 tgaatgcagtaaaacaaagaccgg---ggtagcaattacactcgatttggcgtccgggga 2757
Db 4802 AAGTGTGGGCTCAATCAACATCAAAATGTTGTTAGCTATATCTCGCTACGAGCATCAGGAC 4861
Qy 2758 gaaatttcagtttagcatttgaatgaagttaa 2791
Db 4862 GAAACTATACCTTAACATTAGAATGAATTCCTA 4895
RESULT 11
HIU15058
LOCUS
DEFINITION
Haemophilus influenzae SB33 transferrin binding protein 1 (tbpl)
gene complete cds.
ACCESSION
U15058
VERSION
U15058.1
KEYWORDS
GT:1223954
SOURCE
Haemophilus influenzae.
ORGANISM
Haemophilus influenzae
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Haemophilus.
REFERENCE
1 (bases 1 to 5144)
AUTHORS
Loosmore,S.M., Yang,Y.P., Coleman,D.C., Shortreed,J.M.,
Englund,D.M., Harkness,R.E., Chong,P.S. and Klein,M.H.
Cloning and expression of the Haemophilus influenzae transferrin
receptor genes
Mol. Microbiol. 19 (3), 575-586 (1996)
JOURNAL
MEDLINE
REFERENCE
96228706
2 (bases 1 to 5144)
AUTHORS
Loosmore,S.M.
TITLE
Direct Submission
JOURNAL
Submitted (23-SEP-1994) Sheena M. Loosmore, Connaught Centre for
Biotechnology Research, Molecular Genetics Research, 1755 Steeles
Avenue W., Willowdale, Ontario, M2R 3T4, Canada
FEATURES
source
1. .5144
/organism="Haemophilus influenzae"

/strain="SB33"
/specific_host="Homo sapiens"
/db_xref="taxon:727"
/clone="JB-1031-2-9"
2135..4870
/gene="tbl1"
2135..4870
/gene="tbl1"
/codon_start=1
/transl_table=11
/product="transferrin binding protein 1"
/protein_id="AAC43934.1"
/db_xref="GI:1223955"
/translation="MTKKPIFRISIIISCLISIVVKAETOSIKDKAEISSEVDYOST
EDSLEFISVTAEKIRDKNEVTGLKIIKTSIESREQLNIRIDTRYDPGISVVE
QGRGASGYIRGMDNRNVALVDGLPOTQYVYVPLVARSYSGTGAINIEYENV
KAVEISKSSSGVNGALAGSVTFQSGSAADILEGDKSWGIOITKNAYSKNKGFTHS
LAVAGKGGEGVAIVTQNSERTVHKDALGVQSVREFIATDKSSGFVJQGPCP
NGDDKCAKYPKALSPQSEVTVSDYTGANRIAPNPKMYESQWFLRGGHFSEQHYI
GGIFEFTQQKFDIMFTFPAYLRSRKRDRTGFFPKQDYGAIQRIEDRGVNYASG
LYFDEHHRKQVGEIYIYENKNKAGIIDKAVLSANOQNIILDSYMRHHCGLYPNPSK
NCRPTLDKPYSYRSDRNVIYKERNMLQLNLEKKIQONWLTHQIVNGLPDDTSALQ
HKDYLTRVATANIISGTVAGRRNGYERQPLYISKPVDVFGQDHCNKGSSNYS
DKVRLIKGNKYFAARNNALGKYIDLGIGIRYDVSRTKANESTLSVGFKNFSNT
GIVKPEWLDLSYRLSTGFNPSPFAEMYGRWGNNSDYVYKFPETSRNDFGLA
LKGDGFIENISHSNATRIAFAEELSKNGITGKNGYGHNAONAKLVGNITAOGLD
FNGUKRIPYGTATFAYNRVKVDQKINAGLASVSYLFDIAPQSYRIIGLVHDHPS
NTWGIINTFQSKAKSONELLGQALGNRSNRYKSTRKLTRAWHILIDVSGYYMANKNI
MLRLGIVNLFNYRYVTWEAVROTAQAVNQHVNGSVTRYAASGRNYTLFLEKF"

BASE COUNT 1813 a 888 c 974 g 1469 t

ORIGIN

Query Match 11.98; Score 332.6; DB 2; Length 5144;

Best Local Similarity 50.5%; Pred. No. 9.7e-68;

Matches 1353; Conservative 0; Mismatches 1134; Indels 190; Gaps 15;

QY 173 aaacgaagtaactggcttgggaaagtagtgaataattatcacgaatgagtaaaatc 232
DB 2325 ATAATGAAGTAACGTGGACCTTGGCAAAATTTATAAAACGAGTGAAGTATCAGCGGAGAAC 2384
QY 233 aaattcttggtattcggtattactcgctatgacctgatttctggtgggaacaag 292
DB 2385 AAGTATTAAATTTTCGTGATCTACACGCTATGATCAGGACATTTTCAGTTGTAGAACAAAG 2444
QY 293 gtcgcggtgcgaagtagtggtcattgccattcgaggtgtagataaaacccgtgcagcttac 352
DB 2445 GTCGCGGTGCAAGTCTGGATATCTATTCTGCTGATGGACAGAAATAGAGTTGCTTTAT 2504
QY 353 ttgttgatgggtcaccacaagcgacaggtatcatcagctagggttc----- 398
DB 2505 TAGTAGATGGTTTACCTCAACGCAATCTTATGTAGTGAAGGCCCTTAGTTGCTCGTT 2564
QY 399 ----agatgctaagtggtgccaattaatgagattgagattgataaaacattggtccaattg 454
DB 2565 CAGGATATTTGCGCAGCTGGTGCATTAATGAATTTGAATATGAATAATGTAAGGCCCGTCG 2624
QY 455 agttaagcaagggagcaagttctcggaatatggctctggtgcgcattggtggtctattg 514
DB 2625 AAATAAGCAAGGGGGGNGTCTTCTGAGTATGTTAATGAGCACTAGCTGGTGTCTGTAA 2684
QY 515 gtttcgtactaaagatgycgaggaattatttaaagaggggcagcattgggggcttagata 574
DB 2685 CATTTCAAGCAAAATCCGACGCCGATATCTTTAGAAGGAGACAAATATCATGSGGAATTCAAA 2744
QY 575 gtaagacctttatgcagcaaaaaatagccatttttttacagctctatcgacggcgctggtg 634
DB 2745 CTAATAATGCTTTATTTCAAGCAAAAAATAAGGCTTTTACCCCATTTCTTTAGCTGTAGCAGAA 2804
QY 635 aggcgggtggttttgaagcactgttattgcaactaccacacagcggttaaagagacacaaa 694
DB 2805 AACAAAGTGGNTTGAAGGGGTGCCATTTACACTAACGAAATTCGGAGGAAACCCCAAG 2864

QY 695 ttcatccgagggcaataataataaaacataataattcggcggtataaccggcgttttgaataac 754
DB 2865 TCCATAAAGATGTCATTAAAGCGGTACAAAGTTATTAGCGGATTCATCGCC-----ACRACA 2920
QY 755 gctacgactttaccataataatccgcacagaagtgcctcctggaggatcctcctttaaattgtg 814
DB 2921 GATAAATCTTTCAGGATACCTTTGTGATACAAAGGTGAGTG-----TCCAAATGGTG 2969
QY 815 aagatacttgcacaacacattagattgtactcctcgtgcgaaggttaagttagaacgcgcgata 874
DB 2970 ATGACAAAGTGTGCAGC----- 2985
QY 875 attcccagtgagaacatttcccggaataatacgctgaagagcgcaaacagcgttgagcaga 934
DB 2986 -----CAAACCACCTGCAAGT 3002
QY 935 ttcttatcgactgacgacgtcctcagcccaagaataataccggttaaaagatgcgaattgcac 994
DB 3003 TATCCCCCAAGCGAAACCGTAAGCGTTTCAGATTATACGGGGCTTAACCGTATCAAC 3062
QY 995 caaaccttttagattacaagagtaattctgttttttagaagtttggctatcaacttcaact 1054
DB 3063 CTATCCAAATGAATATGAAGCCAGCTCTGGTTTTTAAGAGNGGATCAITTTTTCTG 3122
QY 1055 cgtctattcttggcgcaatctttagaagatacaaaaacacgcctacgatatcctgata 1114
DB 3123 AACAACTACTATATGGTGTATTTTGAATTCACACAAACAAATTTGATATCGTGATA 3182
QY 1115 tgcaaacgcagcttactatcacaaagacgatatataacttatacttaggaacatatttt 1174
DB 3183 TGACATTTCCCGCTTATTTAAGATCAACAGAAAAACGGGATGATAGAACTGGCCCTTTT 3242
QY 1175 atgaaggggataattatttagatggcttagtttcaagccaaggatcccttaagggttgc 1234
DB 3243 ATCCAAAGCAAGTATATGGTGCATATCAAGTATTAGGAGTGGCCGA-----GGCGTTA 3296
QY 1235 gctatagccatgtgaagtttttttgatgaacgctcaccaacaacgctcgtttaggattcacct 1294
DB 3297 ACTATGCAAGTGGCTTTATTTTCGATGAACACCATAGAAACAGCGTGTAGGTATTGAAT 3356
QY 1295 ataaataaaacacgagaataaactcgtcgtgttgatagcattaaactc---agtgcgata 1351
DB 3357 ATATTTACGAAAAATAAGACAAACGCGGCATCATTTGACAAAGCAGTGTAAAGTCTAATC 3416
QY 1352 acaagatatgaactatagccggtctacatcgctgcattgttagcattagcattatcgtg 1411
DB 3417 ACACAAACATCATCTATGACAGTTATATGACATATGACATATCGTTCTCATAGAA 3476
QY 1412 tagataaaattccgcccgcacttttggaataaattccttggtctatgtatcgaactgagcgt 1471
DB 3477 CAAGTAAGAATTGCGGCCGACACTTGTATAAACCCTTATTTCATCTACTCGTTCTCATAGAA 3536
QY 1472 ataattaccaagaaagcattcgtgctcattcatttagaatttgaataaagcgcctaactgctg 1531
DB 3537 ATGTTTATAAGAAAAACATAATATGTTGCAATTTGAATTTAGAGAAAAAATTCACAAA 3596
QY 1532 gtcaagcgctatttaacccaaaccccaaaactgaattaggg-----ttgggcttgg 1582
DB 3597 ATTGGCTTACTCATCAAAATTTGCTTCAANTCTTGGTTTGTAGTACTTTACTACGCGCTTC 3656
QY 1583 atcgatttaattcgcttatggatcattggggtatgactgctgcccataataaccaagggcgtt 1642
DB 3657 AGCATAAAGATTATTTAACTCGACGCTTACCGCTACGGCAATATATTTATTCAGGGACAG 3716
QY 1643 ataccagctaccgggttagaggcgcttttagataataatccatatatttcgcgcgacccac 1702
DB 3717 TTGCTGTAAACGAAGAAATGGTTAGCAAAAAACACCTTACTTATCTACTCAAAACCAAAAG 3776
QY 1703 gcagttatgaacgggtatcttctgtgtaataataacacgcggc-----gacatcttaact 1756
DB 3777 TAGATTTTGTAGGACAAAGATCATTTGTAATTTATAAGGTAGTCTCTCTTAATACAGCGACT 3836
QY 1757 gtgaaccgcgttaaaataaaggcgatagccattttgttagcttccgc---gatctagtga 1813


```
Db 3837 GTAAAGTCGGGTTAAATTAAAGGGAATAATTATTTTCGCGACGCAATAAATATGCGAT 3896
Qy 1814 taecgagatgtggtgattgggattaggggtgogtttggatcaacatcgatttaaatctg 1873
Db 3897 TAGGAAATACATTGATTTAGGTTTAGGTATTCGGTATCGGTATCTCGTACAAAGCTA 3956
Qy 1874 atgacggtggacacttagccgaacttatcgaaatttggtcttggaaatgggtgggattacgc 1933
Db 3957 ATGAATCAACTATTAGTGTGGTAAATTTAAATAATTTCTCTTGGAAATACGTTATGTC 4016
Qy 1934 ttaacaaacagagttgtatcgctttcttatcgctatccaaacgggttttagagtgctg 1993
Db 4017 TAAACCAACGGATGCTTCATCTTCTTATCGCCTTCTACTGGATTTAGAAATCCTA 4076
Qy 1994 cattctatgaactttatggtaacgtagatcatattgggcttaagataaagaatatgtgc 2053
Db 4077 GTTTTGTGAAATGTATGTTGGCGGTATGGTGCAATAATAGCGATGTTTATGTAGGTA 4136
Qy 2054 aacgcgcgaacgtagccaccagttagagccagaaaaaatcgactaatcatgagattggag 2113
Db 4137 AA-----TTTAAGCCTGAACATCTCGTAAACCAAGAGTTTGGTC 4175
Qy 2114 ttagtcttaaaaggtcgaatttggttacccttgatgtgagctatttccgttaataactataaaa 2173
Db 4176 TCGCTCTAAAAGGGGATTTTGGTAAATATGAGATCAGTCATTTTAGTAAATGCTTATCGAA 4235
Qy 2174 atatga-----ttgcacagcatgtaaaagaataataacaaaaatcacactgtttctata 2227
Db 4236 ATCTTATCGCCTTTGCTGAAGAACTTAGTAAATAATGAACTACTGGAAGGGCAATATG 4295
Qy 2228 actaccataaatcaagatagcactaaacgggataaaatttagtcgtaaaatttgact 2287
Db 4296 GATATCAATAATGCACAAATCAAAATAGTTGGCGTAAATATAACTGCGCAATTAGATT 4355
Qy 2288 tacacggtattttatctatgtgcgaatggtttttatctcatcagttgcttataaacggtg 2347
Db 4356 TTAATGTTTATGGAACGATTTCCCTACGCTTGGTATGCAACATTTGCTTATAACCGAG 4415
Qy 2348 taaagtaaaagagcgaaactaacccgactcaagactcgatagcgtaaacgatcctatcc 2407
Db 4416 TAAAGTTAAAGATC---AAAAATCAATGCTGTTTGGCCCTCGGTAAGCAGATTATTAT 4472
Qy 2408 tagatcgatcagccagcagcatgtgcttggtatcggttcggtacagcaccaccagaagaa 2467
Db 4473 TTGATGCCATTACGCCACGCGGTATATCATTTGTTAGGCTATGATCATCCAAAGTAATA 4532
Qy 2468 aatggggaatttggcattactaccactattctaaagccaaacgcccagtgaggtggcag 2527
Db 4533 CTGGGGAAATTAACATGTTTACTCAATCAAAAGCAAAATCTCAAAATGAATTGCTAG 4592
Qy 2528 gcacacgtcatcagggatcacatcgc-----gttgatttaggtggcaaacctgaccggtt 2580
Db 4593 GACAACGTCATTTGGGTAAACAATTCAAGGAATGTAAATCAACAAGAAAATCTTACTCGGG 4652
Qy 2581 ctgggtacacccatgatattccgg---ttacatcaattataaaaaactacaccttaagt 2637
Db 4653 CATGGCATATCTTAGATGATTCGGGTATTATCATGTTGTTAGGCTATGATCATCCAAAT 4712
Qy 2638 gaggaattataatgtgactaatgttaaatattccacttgggaatcagtcgcccgaatccg 2697
Db 4713 TAGGGATATATAATTTATCAACTATCCCTATGTACTTGGGAAGCGGTGGCTCAACAG 4772
Qy 2698 gtgtgaatcgagtaaaccaagacc---ggggtagcaattacactcgatttggcgtccgg 2754
Db 4773 CACAAGGTGCGGTCAATCAACATCAAAATGTTGGTAGCTATATACCTCGCTACGCAGCATCAG 4832
Qy 2755 ggagaaatttcagtttagcatttgaatgaagtttta 2791
Db 4833 GACGAACATATACCTTAACATTAGAAATGAATTTCTA 4869
```

RESULT 12

```
I79999
LOCUS 5144 bp DNA PAT 10-JUN-1998
DEFINITION Sequence 105 from patent US 5708149.
ACCESSION I79999
VERSION I79999.1 GI:3208289
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5144)
AUTHORS Loomore,S., Harkness,R., Schryvers,A., Chong,P., Gray-Owen,S.,
Yang,Y., Mordin,A. and Klein,M.
TITLE Method for producing purified recombinant Haemophilus influenzae
transferrin binding proteins
JOURNAL Patent: US 5708149-A 105 13-JAN-1998;
FEATURES Location/Qualifiers
source 1..5144
BASE COUNT 1813 a 888 c 974 g 1469 t
ORIGIN
Query Match 11.9%; Score 332.6; DB 5; Length 5144;
Best Local Similarity 50.5%; Pred. No. 9.7e-68;
Matches 1353; Conservative 0; Mismatches 1134; Indels 190; Gaps 15;
Qy 173 aaaaagaagtaactgcttggggaagtagtgaaaaattatcacgaaatgagtaaaaaatc 232
Db 2325 ATAATGAAGTAACGTGACTTGGCAAAATATATAAAACGAGTGAAGTATCAGCCGAGAAC 2384
Qy 233 aaattcttggtattcggtatttaactcgctatgacctggtattctcggtggtggacaag 292
Db 2385 AAGTATTAAATATCGTGATCTAACAGGCTATGATCAGGCAATTCAGTTGTAGACACAG 2444
Qy 293 gtgcggtgcgaagtagtggctatgccattcgaggtgtagataaaaaacccgtgtcagcttac 352
Db 2445 GTCGCGTGCAGTTCCTGGGATATTTCTTATCGTGTATGGACAGAAATAGAGTTGCTTTAT 2504
Qy 353 ttgttgatgggtcacacaagcgcaggtatcatcagctaggttc----- 398
Db 2505 TAGTAGATGTTTACCTCAAAACCAATCTTATAGTGTCAAAAGCCCTTTAGTTGCTCGTT 2564
Qy 399 ----agatgctaattggtgtgcaatgaatgagattgagatgaataaacattcgttcaattg 454
Db 2565 CAGGATATTCGCACTGGTGCATTAATGAATTTGAATGAATGAATGAATGAATGAATGAAT 2624
Qy 455 agttaagcaagagagaagttctgcggaatatgctctggtgcgcgcatggtggtgctattg 514
Db 2625 AAATAAGCAAGGGGGGAGTTCCTCTGAGTATGTTAATGGAGCACACTAGTGTGTTCTGTAA 2684
Qy 515 gtttcgtactaaagatgcgcagagatattataaagaggggcagcattggggcttagata 574
Db 2685 CATTTCAAGCAAAATCCGCGAGCGATATCTTAGAAGGAGACAAATCATATGGGAATTCAAA 2744
Qy 575 gtaagacctttatgcgcagcaaaaatgccatttttacagttctatccgcgcgctggtg 634
Db 2745 CTAAATATGCTTATTCGAAGCAAAATAAAGGCTTTACCCATTCCTTTAGCTGTAGCAGGAA 2804
Qy 635 aggcgggtgttttgaagcactgtttattgcaactaccgcacacggttaaagagagaccataa 694
Db 2805 AACAAAGTGGATTGAAGGGTGCCTCATTTACACTCAACGAAATTCGGAGAGAAACCCCAAG 2864
Qy 695 ttcattccgagggaataataataacataaatatccgcgtataaacccggtttgaaaaatc 754
Db 2865 TCCATAAAGATGCAATTAAGGGGTACAAAAGTTATAGCGGATTCATCGCC-----ACAACA 2920
Qy 755 gctacgactttacccaattccgcagagaatgctccctggaggatctctctttaaattg 814
Db 2921 GATAAATCTTCAGGATACTTTGTGATACAAAGGTAGTG-----TCCAATAGGTG 2969
Qy 815 aagatacttgcccaacattagattgtactctcgtgcgaagggttaagttgaacgcgcgata 874
Db 2970 ATGACAAGTGTGCAGC----- 2985
```


Db 3081 TGACATTCCCGCTATTATTAAGGCCACAGAGGATTTACAAAGTGCGCCCTTTT 3140
QY 1175 atgaagggaataatttttagatgcttagtggttcaagccaagatcccttatgggtgc 1234
Db 3141 ATCCAAAGCAGAT- ----TATGGTCATATCAACATATTGGTGATGGCAGAGCGCTTA 3194
QY 1235 gctatagccatgtgaagttttttgatgaacgtaccacaacgcgtgttaggattcacct 1294
Db 3195 AATATGCAAGTGGCTTATTTTCGATGAACACCATAGAAAACACGCGTGTAGGTATTGAAT 3254
QY 1295 ataaataaaccagagaataa ---tcgctgggttggtatagcattaaactcagtcgagata 1351
Db 3254 ATATTTACGAAATTAAGAACAAAGCGGCATCATTTGACAAAGCGGTGTTAAGTGCTAATC 3314
QY 1352 aacaagatttgaactatatagcgcgtacatcgcttgcatgttgatgagcattacctgtg 1411
Db 3315 AACAAACATCATACTTGCAGATTATATGCGACATACGCATTCGAGCTTTATTCCTAATC 3374
QY 1412 tagataaaattgcgcgcgcactttggataaatttggctatgtatcgaactgagcgta 1471
Db 3375 CAAGTAAGAATTGCGGCCAACACTTGATAAACCTTATTCATACATATCATCTTGATAGAA 3434
QY 1472 ataattaccaagaagcatcgtctcattcatttagaatttataaagcgtcaaatgctg 1531
Db 3435 ATGTTTATAAGAAAAACATAACAGGTTCGAATTTGAATTTAGAGAAAAAATTCAACAAA 3494
QY 1532 gtcaaggcgtatttaacccaaacccacacactgaatttaggttggttgcttgatcgatta 1591
Db 3495 ATTGGCTTACTCATCAAAATGTCCTTCAATCTTGCTTTGA ---TGACTTTACTTCCGCAC 3551
QY 1592 attcgttatggatcatcgtgggatagactgcctcccaataaccgaaggcgttataccagct 1651
Db 3552 TTCAGCATAAAGATTATTAACTCGAGCTGTATCGCTACGGCAAGTAGTATTTCAGAGA 3611
QY 1652 acgcggtagaggcgtttagat-----aatccatatatttatcgcgcgac 1699
Db 3612 AACGTGTGAAGCAAGAAAGAAATGGTTTACAATCAAGCTCTTACTTATACCCACACCAA 3671
QY 1700 cagcagattgaacggtatcttctgtgtaataacag-----cgcgacatcttaa 1753
Db 3672 AAGCAGAGTTGGTAGGAGAGACTTTGTAAATATCAAGTAAGTCTCTAAATACAGTG 3731
QY 1754 actgtgaaccgcgtaaattaaagcgcatagccattttgttagcttcgc---gatctag 1810
Db 3732 ACTGTAAAGTCGGTTAATTAAGGAAAAATTTATTTCGCAGCACGCAATAATATGG 3791
QY 1811 tgataagcgatgtgtgatttgggattaggggtcggttttgatcaacatcgatttaaat 1870
Db 3792 CATTAGGGAATACGTTGATTTAGGTTTAGGTATGAGGTATGAGGTATCGTATCTCGTACAAAAG 3851
QY 1871 ctgatgacgcgtgacacttagccgaacttatgaaattggtctggaatggtgggatta 1930
Db 3852 CTAAATGAATCAACTATTAGTGTGGTAAATTTAAATAATTTCTCTGGAAATCTGGTATTG 3911
QY 1931 cgcttaaaccaacagagtttggatcgctttcttcatcgatttcaaacggttttagatgc 1990
Db 3912 TCATAAACCAACAGGAATGCTTGTATCTTCTATCGCTTCTTCTAGTGATTTAGAAATC 3971
QY 1991 ctgcattctatgaactttatggttaaacgctgatcataatgggctttaagataaagaatag 2050
Db 3972 CTAGTTTGTGAAATGATGTTGGGTTGGGTATGGTGAAGGATACCGATGTTTA----- 4026
QY 2051 tgcaacgcgcgaacgtagccaccagtagcagagaaataatgactaatcatgagattg 2110
Db 4027 -----TATAGGTAAATTTAAGCCTGAAACATCTCGTAACCAAGAGTTTG 4070
QY 2111 gagtagctttaaggttcaatttgggttacccttgatgtgagctatttccgtaataactata 2170
Db 4071 GTCTCGCTTAARAGGGATTTGGTAAATTTAGATCACTCATTTTAGTAATGCTTATC 4130
QY 2171 aaatatga-----ttgcagacgcatgtaaaagaataatacaaaaatcacactgttct 2224

Db 4131 GAAATCTTATCGCCTTTGCTGAAGAACTTAGTAAAAATGGAAGTACTTACGGAAGGCAATT 4190
QY 2225 ataactaccataatattcaagatgtagcactaaacgggataaatttagtcgctaaattg 2284
Db 4191 ATGGATATCATATATGACAAAATGCAAAATAGTTGGCGTAAATATATACCTGCGCAATTAG 4250
QY 2285 acttaccggtattttatctctatcgtccagatggttttttattcatcagttgctttataacc 2344
Db 4251 ATTTTAATGTTTATGAAACGCTATTCCTACGTTGGTATGCAACATTTGCTTATTAACC 4310
QY 2345 gtgtaaaagttaaaagcggaactaaacgactcaagactcgatagcgtaaacatccta 2404
Db 4311 GAGTAAAGTTAAAGATC-----AAAAATCAATGCTGTTTAGCTTCGCTAAGCAGTTATT 4367
QY 2405 ttctagatcgattcagcagcagcgtatgcttgcattcgcgtacgctacgacacccagaag 2464
Db 4368 TATTTGATGCCATTCAGCCAGCGGTATATCATTTGTTTAGGCTATGATCATCCAAAGTA 4427
QY 2465 aaaaatggggaattggcattaccacacctattctaaagccaaacgcgcgtaggtg 2524
Db 4428 ATACTTGGGAATTAAGACAAATGTTTACTCAATCAAAAGCAAAATCTCAAAATGAATTGC 4487
QY 2525 caggcacacgtcatcagcgatcacatgc-----gttgatttaggtggcaaacgtgacg 2577
Db 4488 TAGGAAAACGTGATTTGGGTAACAATTCAGGAATGTAAATCAACAAGAAATTTACTC 4547
QY 2578 gtcttggtacacccatgatattaccggt---ttacatcaattataaaacacacacctac 2634
Db 4548 GGGCATGCCATATCTTAGATGTTATCGGTTTATTCATGTTGATGAATGAAGTATTTTGTTC 4607
QY 2635 gtggagaatttataatgtgactaaatcgtaaatattccacttgggaatcgatgcgcgaat 2694
Db 4608 GATTAGGAGTATATAATTTTAACTATCGTATGCTATGCACTTGGGAAGCGTGGTCAAA 4667
QY 2695 cgggtgtgaatgagtaaaccaagacc---gggtagcaattacacactcgatttggcgctc 2751
Db 4668 CAGCACAAAGTCGGTCAATCAACATCAAAATGTTGTAATCTATCTATCTCGCTACGAGCAT 4727
QY 2752 cggggagaatttcagtttagcatttgaatgaatttta 2791
Db 4728 CAGGACGAAACTATACCTTAACATTAGAAATGAAATTTCTA 4767

RESULT 14

LOCUS HIU15052 5009 bp DNA BCT 02-APR-1996
DEFINITION Haemophilus influenzae Minna transferrin binding protein 1 (tbp1) and transferrin binding protein 2 (tbp2) genes, complete cds.
ACCESSION U15052
VERSION U15052.1 GI:1223940
KEYWORDS Haemophilus influenzae.
SOURCE Haemophilus influenzae.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Haemophilus.
REFERENCE 1 (bases 1 to 5009)
AUTHORS Loosmore, S.M., Yang, Y.P., Coleman, D.C., Shortreed, J.M., England, D.M., Harkness, R.E., Chong, P.S. and Klein, M.H.
TITLE Cloning and expression of the Haemophilus influenzae transferrin receptor genes
JOURNAL Mol. Microbiol. 19 (3), 575-586 (1996)
MEDLINE 96228706
AUTHORS Loosmore, S.M.
REFERENCE 2 (bases 1 to 5009)
AUTHORS Loosmore, S.M.
TITLE Direct Submission
JOURNAL Submitted (23-SEP-1994) Sheena M. Loosmore, Connaught Centre for Biotechnology Research, Molecular Genetics Research, 1755 Steeles Avenue W., Willowdale, Ontario, M2R 3T4, Canada
FEATURES
Location/Qualifiers
1. 5009
/organism="Haemophilus influenzae"
/strain="Minna"
/specific_host="Homo sapiens"

[illegible]

```
Db 3519 ATGTTTATAAGAAAAACATAATATGTTGCAATTGCAATTTAGAGAAAAAATTCACAAA 3578
Qy 1519 --gogctaaatgctgtgcaagcgctatttaacaaacccaaacaaactgaatttaggttgg 1576
Db 3579 ATTGGCTTACTCATCAAAATGCTCTCAATCTTTGTTGTTTGGTACTTACTTCAGCGCTC 3638
Qy 1577 gctttgatcgatttaattcogcttagtgatcgatcggtggtatgactgcccaataaccagaag 1636
Db 3639 AGCATAAAGATTAATTAATCTGCACGTGTATCGCTACGGCAGATAGTATTCGAAGGAAAC 3698
Qy 1637 gcgggttaccagctaccgcggtgagggcggttttagataaattccatattatttgcgcgcg 1696
Db 3699 CTGGTGAACCTGGTAAACCAAGAAATGTTGCCATCAACACCTTACTTATACCCAAAC 3758
Qy 1697 atccacgcagttattgaacggtgatctttgtgtaataatacacgcgcg-----gacatct 1750
Db 3759 CAGAGCCATATTTTCAGGACAAAGATCATTTGTAATTTATCAAGGTAGCTCCTCTAATTACA 3818
Qy 1751 taaactgtgaaccggtgtaaaattaaagcgatagccattttgttagcttcgcg---gatc 1807
Db 3819 GAGACTGTAAAGTGGGTAAATTAAGGGAATAATTTATTTTCGCACGACGCAATAATA 3878
Qy 1808 tagtgataagcgagtgatgtgatttggttaggggtgctgttttgatcaacatcgattta 1867
Db 3879 TGGCATTTAGGGAATACGTTGATTTAGGTTTAGGTATTTCGGTATGACGTATCTCGTACAA 3938
Qy 1868 aatctgatgctcgctggacacttagccgaacttatcgaaatttggcttgggaatgggtgga 1927
Db 3939 AAGCTAATGAATCAACTATTAGTGTGTGTAATTAAGGGAATAATTTCTTTGGAATACTGCTA 3998
Qy 1928 ttacgttaaaccaacagagttgtgatcgctttcttatttcgatccgaattcaaacggttttagag 1987
Db 3999 TTGTCATAAAACCAACGGAATGGCTTGATCTTCTTATCGCTTTTCTACTGGATTAGAA 4058
Qy 1988 tgcctgcattctatgaactttatgtataacggtgatacatatggctttaaagataaacgaat 2047
Db 4059 ATCCTAGTTTTCTCAAAATGATGTTGCGGTATGTTGGCGAAGTAACGAGGTTTATG 4118
Qy 2048 atgtgaacgcgcgaacgtagccaccagttagagccgaaataatcgactaatcatgaga 2107
Db 4119 TAGGTAAA-----TTTAAGCCTGAAACATCTCGTAACCAAGAGT 4157
Qy 2108 ttggagttagcttttaaggttcaattgtttacccttgatgtgagctatttcogtataaact 2167
Db 4158 TTGGTCTCGCTCTAAAGGGGATTTTGGTAATATTAGATCAGTCATGTCATTTAGTAATGCTT 4217
Qy 2168 ataaaaatagattgcagacgatgaaa---agaataacaaaaatacacactgtttct 2224
Db 4218 ATCGAAATCTTATCGCCTTTGCTGAAGAACTTAGTAAAAAATGGAACCTGGAAGGGCAAT 4277
Qy 2225 ataaactacaataattcaagatgtagcactaaacgggataaaatttagtcgctaatttg 2284
Db 4278 ATGGATATCATATGCACAAAATGCAAAATTAGTTGGCGTAAATATATACTGCACANTTAG 4337
Qy 2285 acttacacggttatttatactatgctgcagatggtttttatttcattcatcagttgottataacc 2344
Db 4338 ATTTTAATGGTTTATGGAAAGCATATCCCTACGGTTGGTATGCAACATTTGCTTATAACC 4397
Qy 2345 gtgtaaaaagtaaaagcggaactaaacgcgactcaagctcgatagcgtaaacgatccta 2404
Db 4398 AAGTAAAAAGTTAAAGATC---AAAAATCAATGCTGGTTTAGCCCTCCAGCAGTTAT 4454
Qy 2405 ttctagatgcattcagccagcacgctatgtgcttggttgcgttcgctacgatcacccagaag 2464
Db 4455 TATTGTATGCCATTCAGCCACGCCGTTATATCATTTGGTTTAGGCTATGATCATCCAGTA 4514
Qy 2465 aaaaatggggaattggcattactaccactattcttaagcagcaaaaccccgatgaggtgg 2524
Db 4515 ATACTTGGGAATTAATACAATGTTTACTCAATCAAAAAGCAAAATCTCAAAATGAATTCG 4574
Qy 2525 caggcacagtcatacagatatacatcgc-----gttgatttaggtggcaaacactgacgc 2577
Db 4575 TAGGAAAACGTCATTTAGGTAAACAAATTCAGGGATGTAATAATCAACAAGAAAACCTTACTC 4634
```

```
Qy 2578 gtttttggtacacccatgatattaccgg---ttacatcaattataaaaaactacaccttac 2634
Db 4635 GGGCATGTCATCTTAGATGTATCGGTTATTACATGCGAATAAAAAATATTATGCTTC 4694
Qy 2635 gtgaggaattataatgtgactaatcgtaaatattccacttgggaatcagtgcgccaat 2694
Db 4695 GATTAGGATATAAATTTTATTCACATATCGCTATGTTTACATGGAACGGGTGCGTCAAA 4754
Qy 2695 ccggtgtgaatgcagtaaaaccaagaccg---gggtagaattacactcgatttggcgctc 2751
Db 4755 CAGCACAGGTGCGGTCAATCAACATCAAAATGTTGGTAGCTATACTCGCTACGAGCAT 4814
Qy 2752 cgggggagaaatttcagtttagcatttgaataagaagtttta 2791
Db 4815 CAGGACGAACCTATACCTTAACTTAGAATGAATTTCTA 4854
```

RESULT 15

```
LOCUS I79989 5009 bp DNA PAT 10-JUN-1998
DEFINITION Sequence 3 from patent US 5708149.
ACCESSION I79989
VERSION I79989.1 GI:3208279
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 5009)
AUTHORS Loosmore,S., Harkness,R., Schryvers,A., Chong,P., Gray-Owen,S.,
Yang,Y., Murdin,A. and Klein,M.
TITLE Method for producing purified recombinant Haemophilus influenzae
transferrin binding proteins
JOURNAL Patent: US 5708149-A 3 13-JAN-1998;
FEATURES Location/Qualifiers
source 1..5009
BASE COUNT 1809 a 836 c 934 g 1430 t
ORIGIN
```

Query Match 11.98; Score 331.2; DB 5; Length 5009;
Best Local Similarity 50.5%; Pred. No. 2.1e-67;
Matches 1354; Conservative 0; Mismatches 1133; Indels 193; Gaps 15;

```
Qy 173 aaaaagagtaactggcttgggaaagtagtaaaaattatcacgaaatgataaaatc 232
Db 2307 ATAATGAAGTAACCTGGACTTGGCAAAATATCAAAACTAGTGAAGTATACGCCGAGAAC 2366
Qy 233 aaattcttggtattcgtgatttaactcgctatgacctgggtatttcggtggtgaacaag 292
Db 2367 AAGTATTAAATATTCGTGATCTAACACGCTATGATCCAGGGATTTCACTTGTAGAACAAG 2426
Qy 293 gtccggtgcgaagtagtggctatccattccaggtgtagataaaaaacccgtgtcagcttac 352
Db 2427 GTCGGGTGCAAGTCTTGGATATTCTATTCGTGTTATGATGACAGAAATAGAGTTGCTTTAT 2486
Qy 353 ttgttgatgggtaccacaagcgacagattatcatacgttaggttc----- 398
Db 2487 TAGTAGATGGTTTACCTCAAAACGCAATCTTATGTAGTCAAAAGCCCTTTAGTTGCTCCTT 2546
Qy 399 ----agatgctcaatgggtgctcaattaatgattgagattgagtaaaacattcgttcaattg 454
Db 2547 CAGGATATTCGCGACCTGGTGCATTAATGAATTAATGAATGAATAATGAAGCCGCTCG 2606
Qy 455 agttaagcaagggaggaagttctcgcgaatattggtctcgtgcgcatggtggtgctattg 514
Db 2607 AAATAAGCAAGGGGGGGAGTCTTCTGAGTATGTTAATGGAGCACTAGCTGTTCTGTAA 2666
Qy 515 gtttcgtactaaagatgcgcagatattataaagggggcgacgattggggccttagata 574
Db 2667 CATTTCAAAGCAAAATCAGCAGCCGATATCTTAGAAGGAGAGACAAATCATATGGGGGAATTCAAA 2726
```

Qy 575 gtaagaccttattgccagcaaaaatagccatttttttacagctctatocgacggctggtg 634
Db 2727 CTAAAAATGCTTATTCAAGCAAAATAAAGCGCTTACCACATCTTTAGCTGACAGGAA 2786
Qy 635 aggcgggtggttttgaagcaactgtttatgcaactaccgcagcaggttaagagagcaaaa 694
Db 2787 AACAAAGTGGGATTTGAAGGGTAGCCATTACACTCAACGAAATTCATTGAAACCCCAAG 2846
Qy 695 ttcttcgcagggcaataataataaacataatttcgcgctataaccggctttgaaatc 754
Db 2847 TCCATAAAGATGCAATTAAGCGGTACAAAGTTATGATCGATTATCGCC-----ACAACA 2902
Qy 755 gctacgactttaccacaattccgcacagaatgctcctcgaggaatcctctttaaattggtg 814
Db 2903 GATAAATCTTCAGGATACCTTTGTGATACAAAGGTGAGTG-----TCCAAATGGTG 2951
Qy 815 aagatacttgcaccaacattagattgtactcctcgtgcgaagggttaagttgaaccgcgata 874
Db 2952 ATGACAAGTGTGCAGC----- 2967
Qy 875 atttcccgagtgaacaacatttcgggaatatagcgcctgaagagcgcaaacagattgagcaga 934
Db 2968 -----CAAGCCACCTGCGACTT 2984
Qy 935 ttccttatcgcactgagcagctctcagcccagaataatcccggtaaagatcgcatgtgcac 994
Db 2985 TATCACCCCAAGCAACCGTAAAGCGTTTCAGATTATACGGGGCTTAACCGTATCAAAC 3044
Qy 995 caaaccttttagattacaagagtaattcgtttttttatgaagtttggctatcaactcaact 1054
Db 3045 CTAATCCAATGAATATGAAAGCCAGTCTTGTTTAAAGAGGGGTATCATTTTCTG 3104
Qy 1055 cgtctctattcttgcgcaactcttagaagatacaaaaaacgcgtacgatatcogtgata 1114
Db 3105 AACAACTATTATTGGTGGTATTTTGAATTCACACACAAATAATTTGATATCCGTGATA 3164
Qy 1115 tgaacgcgcagcttactatatacaaaagcagatattaaacttatcacttaggaactgttt 1174
Db 3165 TGACATTTCCCGCTATTATTAGCCCAACAGAAAGCGGGATGATAGTAGTCGTCTTTT 3224
Qy 1175 atgaagggtataatttttagatggcttagtgttcaagccaagatcccttatgggttgc 1234
Db 3225 ATCCAATGCAAGATCATGGTGATATCAACATATTGAGGATGGCA-----GAGGCGGTA 3278
Qy 1235 gctatagcattgtaagttttttgatgaagtcaccacaacacgctggttaggattcaact 1294
Db 3279 AATATGCAAGTGGGCTTTATTTCGATGAACACCATAGAAAAACAGCGTAGGTATTGAAT 3338
Qy 1295 ataaataaaacccagagataatcgtggttggatagcattaaactc---agtcgggata 1351
Db 3339 ATATTTACGAAAAATAAGAACAAAGCGGGCATCAITGACAAGACAGTGTAAAGTGTAAATC 3398
Qy 1352 acaagatttgaactatagccgggtacatcgttgcattgttagcagattatcctgtg 1411
Db 3399 AACAAAACATCATCTTGACAGTTATATGCGACATACGCATTGCGAGTCTTTATCCTCAATC 3458
Qy 1412 tagataaaaaattgcgcgcgactttgggataaaacttggctctatgatacgaactgagcgt 1471
Db 3459 CAAGTAAAGATTGCCGCCAACACATGTAATAACCTTATTCATCTATCTCTGTATAGAA 3518
Qy 1472 ataatcaagaagaagcatcgtgctcatctcatttagaatttgataaa----- 1518
Db 3519 ATGTTTAAAGAAAAACATATATGTTGCAATTTGAATTTAGAGAAAAAATAATTCACAAA 3578
Qy 1519 --gcgctaaatgctggtcaagcgttatttaaccacaaccacaactgaatttaggttgg 1576
Db 3579 ATTGGCTTACTCATCAAAATGCTTCAATCTGTTTGTGATGACTTTACTTCAGGCGTTC 3638
Qy 1577 gctttgacgatttaattcgtcttagatcatggtgggatatgactgcccaataataccaaag 1636
Db 3639 AGCATAAAGATTATTAATACTCGAGTGTATGCTGACGGCAGATAGTATTCAGAGGAAAC 3698
Qy 1637 ggggttaaccagctaccgcggttagaggcgttttagataaattccatatttatcgcgcg 1696

Db 3699 CTGCTGAACCTGCTAAACCAAGAAATGTTTGCATCACAACCTTACTTATATACCAAAAC 3758
Qy 1697 atccacgcagatttgaacacggtatcttttgtataataatacacgcggc-----gacatct 1750
Db 3759 CAGAGCCATATTTTTCAGGACCAAGATCATTTAATTAAGGTAGCTCTCTTAATACA 3818
Qy 1751 taaactgtaaccgcgtataaattaaagcagtagcgaattttttagcttcgcg---gactc 1807
Db 3819 GAGACTGTAAAGTCGGTTTAATTAAGGGAATAATTTATTCGACAGCGCAATATA 3878
Qy 1808 tagtgaagaagcagattgtgattgggattaggggtgcgttttggatcaacatcgattta 1867
Db 3879 TGGCATTAGGGAATACGTTGATTAGTTAGTATTCGGTATGACGTATCTCGTACAA 3938
Qy 1868 aatcgtatgacgcgtgagacacttagccgaacttatcgaaatttggcttcttgaatggtgga 1927
Db 3939 AAGCTAATGAATCAACTATTAGTGTGGTAAATTTAAAAATTTCTCTTGGAAATCTGGTA 3998
Qy 1928 ttacgcctaaacacagagtttgcgtcttcttctcattcgcatttcaaacggttttagag 1987
Db 3999 TTGTCATAAAACCAACCGAATGGCTTGATCTTTCTTATCGCTTTCTACTGGATTAGAA 4058
Qy 1988 tgcctgcattctatgaactttatgtaaacggtgatcatattgggctttaaagataaacgaat 2047
Db 4059 ATCCTAGTTTTCGAAATGTATGTTGGCGGTATGTTGGCAAGATGACGAGGTTTATG 4118
Qy 2048 atgtgaacgcgcgcacgtagccaccagtttagagccagagaaaaatcgactaatcatgaga 2107
Db 4119 TAGTAA-----TTTAAAGCCCTGAAACATCTCGTAACCAAGAGT 4157
Qy 2108 ttggagttagctttaaaggccaatttggttaccttgcgttgatggagctatttccgtataaact 2167
Db 4158 TTGCTCGCTCTAAAAGGGGATTTTGGTAAATTTGAGATCAGTCATTTTATGTAATGCTT 4217
Qy 2168 ataaaaatgattgcgacagcatgtaaa---agaaataacaaaaaatcacactgtttct 2224
Db 4218 ATCGAATCTTATCGCTCTTGTCTGAAGNACTTACTAANAATGGAACCTGGAAGGCAATT 4277
Qy 2225 ataactaccataatattcaagattgtagcactaaacgggataaaatttagtgcataaattg 2284
Db 4278 ATGGATATCAATAATGCAAAAAATAGTTGGCGTAAATATAACTGCACAAATAG 4337
Qy 2285 acttaacgcgtattttatctatcgtgcagattggttttttattcattcagttgcttataacc 2344
Db 4338 ATTTAATGGTTTATGAAACGTTATCCCTACGGTTGGTATGCAACATTTGCTTATATACC 4397
Qy 2345 gtgtaaaagtaaaagagcggaaactaaaccgactcaagactcgatagcgtaaacgactccta 2404
Db 4398 AAGTAAAGTTAAAGATC---AAAAATCAATGCTGGTTTAGCCCTCGTAAGCAGTTATT 4454
Qy 2405 ttctagatgcattaccgcagcagcgtatgctgtgattgcgtgattgcgtacgataccaccaga 2464
Db 4455 TATTGTATGCCATTACAGCCAGCGTTATATCATTTGTTTATGCTATGATCATCCAACTA 4514
Qy 2465 aaaaatgggaattggcattactactaccacttctaaagccaaacgacgcataggtg 2524
Db 4515 ATACTTGGGGAATTAATACAAATGTTTACTCAATCAAAAGCAAAATCTCAAAATGAATTGC 4574
Qy 2525 caggcaacgcgtatcaacggatatacatcgc-----gttgatttaggtggcaaacgacgc 2577
Db 4575 TAGGAAAACGTGATAGGTAAACAATTAAGGGATGTAANAATCAACAGAAAACTTACTC 4634
Qy 2578 gttcttggtaacccatgatattaccgcg---ttacataaatataaaaaactaccacttac 2634
Db 4635 GGGATGCTATCTTATAGATGATCGGTTTATACATGGCGAATAAAAAATTTATGCTTC 4694
Qy 2635 gtgaggaattataatgtgactaatgtaaatattccacttgggaactgagcgcaat 2694
Db 4695 GATTAGGATATATAATTTATTAACATATCCGTATGTTACTTGGGAACGGGTGCTCAA 4754
Qy 2695 ccggtgtgaatgcagtaaaaccaagacgcg---gggtagcaattatacactcgatttggcgtc 2751

Db 4755 CAGCACAAAGGTGGGTCAATCAACATCAAAATGTTGGTAGCTATACTCGCTACGCAGCAT 4814

Qy 2752 cggggggaatttcagtttagcatttgaagtgaagtta 2791

Db 4815 CAGGACGAACTATACCTTAACATTAGAAATGAAATCTTA 4854

Search completed: September 14, 2000, 00:15:09
Job time: 3752 sec

THIS PAGE BLANK (USPTO)

Query Match 99.18; Score 2766; DB 1; Length 2784;									
Best Local Similarity 99.78; Pred. No. 0;									
Matches 2784; Conservative 0; Mismatches 0; Indels 8; Gaps 1;									
Qy	1	atgataaagaatatcatcttccgtctatccactgttgcccttaacagtggtatttgct	60						
Db	1	ATGATAATGAATATATCATCAATTTTCGTATTCACCTGTGTGCTTAACAGTGTATTGCT	60						
Qy	61	ctttctcatcactcaggtgctgcactgaaataaaaaaaacgaagaaaaataacgatcta	120						
Db	61	CTTTCCTATTATACGGTGCCTGCGACTGAAATAAAAAATCGAAGAAAATAAGCATCTA	120						
Qy	121	gctgtctcagatgaagtatttgaacagagaccattatgctcagacgctcaaaacgaa	180						
Db	121	GCTGTCCTGAGATGAAGTTATTGTGACAGAGAGCCATTATGCTCACGAACGTCAAAACGAA	180						
Qy	181	gtaactggcttgggaaagtagtgaataatcaccgaaatagtgtaaaaaatacaattctt	240						
Db	181	GTAAGTGGCTGGGAAAGTAGTGAATAATATCACGAATAGTAAATAATCAAAATCTT	240						
Qy	241	ggattcgtgatttaactcgtatgacctgtgtatttctgggtggaaacaggtcgcggt	300						
Db	241	GGTATTCGTGATTTAACTCGCTATGACCTGTGTTTCGGTGGTGAACAAGGTGCGGGT	300						
Qy	301	gcaagtagtgctatgccattcgaggtgtagataaaaaccgtgtcagcttacttctgat	360						
Db	301	GCAGTAGTGCTATGCCATTTCGAGGTGTAGATAAAACCCTGTACAGCTTACTTGTGAT	360						
Qy	361	gggtaccacaagcgacagttatcatacgcgtaggttcagatgctaatgggtggcgcaatt	420						
Db	361	GGGTACCACAAGCGCACAGTTATCATACGCTAGGTTCCAGATGCTAATGGTGGTGCATTT	420						
Qy	421	aatgagatgagatgaaaaaattcgttcaattagtttaagtaacaaaggaggaagttctgcg	480						
Db	421	AATGAGATGAGTATGAATAACATTGTTCAATTGAGTTAAGCAAAAGGAGCAAGTTCTGCG	480						
Qy	481	gaatatgctctggcgcatgggtgctctattggtttctactaaagatgcgcagat	540						
Db	481	GAATATGCTCTGGTGCCGANTGGTGCTCTATGTTTTCTGCTACTAAAGATGCCGAGAT	540						
Qy	541	attattaagaggggcagcattggggcttagatagtaagacctcttatggcagcaaaaat	600						
Db	541	ATTATTAAAGAGGGCAGCATTGGGGCTTAGATAGTAAGACCTCTTATGCCAGCAAAAAT	600						
Qy	601	agccaatttttaacgtctatcgcagcgctggtgagggcggtggttttgaagcacttgct	660						
Db	601	AGCCATTTTTTACAGTCTATCGACGGCTGTGAGCGGGTGGTGTTTGAAGCACTTGT	660						
Qy	661	attgcaactcaccgacacggttaaagagaccaaaattcattccgagggcaataaaat	720						
Db	661	ATTGCAACTACCGACACGGTAAAGAGACCAAAATTCATTCCGAGGCAANATAAATTAANA	720						
Qy	721	cataatattcggcgtataaccggctttgaaatcgtctacgactttaccocaaattccgcac	780						
Db	721	CATAATATTGGCGGTATAACGGCTTTGAAATTCGTACGACTTTACCCAAATTCGCGAC	780						
Qy	781	agaaatcctctggaaggaatctccttttaattgtggaagatacttgcccaacattagattgt	840						
Db	781	AGAATGCTCTCTGGAGGATCTCCTTTAATTGTGGAAGATACTTGCCCAACATTAGATTGT	840						
Qy	841	actcctcgtgcaagggttaagtgaaccgcgcatgaatttcccagtgagaaatttccggaa	900						
Db	841	ACTCCTCGTGCAAGGGTTAAGTTGAACCGCGATAATTTCCAGTGAGAACATTTCGGGAA	900						
Qy	901	tatacgcctgaagagcgcaaacagcttgagcagattctcttatcgcactgagcagctctca	960						
Db	901	TATACGCTGAAGACGCAACAGCTTGAGCAGATTCTCTTATCGCACATGAGCAGCTCTCA	960						
Qy	961	gcccaagaataataccggtaaagatgcatttgcaccaaacctctttagatacaaaagagtaat	1020						
Db	961	GCCCAAGAATAATACCGGTAAGATCGCATTGCAACCAAAACCCCTTTAGATTACAAGAGTAAT	1020						
Qy	1021	ctctgttttttagaagtttggcttatcacttcaactcgtctcattatcttggcgcaattcta	1080						

Db	1021	TCGTGTTTTATGAAGTTTGGCTATCACCTTCAACTCGTCTCATATATCTTGGCGCAATCTTA	1080						
Qy	1081	gaagatacaaaaacacgctacgatatccgtgatcatgcaaacgccagcttactatacaaaa	1140						
Db	1081	GAAGATACAAAAACACGCTACGATATCCGTATATGATGCAACGCCAGCTTACTATACAAA	1140						
Qy	1141	gacgatattaacttatacacttaggaactatggtttatgaaggggataatttttagatggc	1200						
Db	1141	GACGATATTAACTTATCACCTTAGGAATATGTTTATGAAGGGGATAATATTTTAGATGGC	1200						
Qy	1201	ttagtgttcaagccaagatcccttattgggttgcgtatagcccatgtgaagttttttgat	1260						
Db	1201	TTAGTGTCAAGCCCAAGATCCCTTATGGGTGGCTATAGCCATGTGAAGTTTTTGTAT	1260						
Qy	1261	gaacgtcaccacaacacgtctgttaggattcacctataataataaacccagagaaataatcgc	1320						
Db	1261	GAACGTCAACCAAAACGTCGTTTAGGATTACCTTATAATATAAAACCCAGAGAATAATCGC	1320						
Qy	1321	tggttgatagcattaaactcagtcggaataaaacaaagatatattgaactatatagcggcta	1380						
Db	1321	TGGTTGATAGCATTAACCTCAGTGGGATAAACAAGATATTGAACATATATAGCCGGCTA	1380						
Qy	1381	catcgtctgatttagcgattatccctgtgtagataaaaaattgccccgacttttgat	1440						
Db	1381	CATCGCTTGCAATGCTAGGATTTATCCTGTGTAGATAAATAATTCGCCGCCGACTTGGAT	1440						
Qy	1441	aaatcttgggtctatgtatcgaactgagcgttaataattaccocaaagaaacacgtgctcatt	1500						
Db	1441	AAATCTTGGTCTATGTATCGAACTGAGCGTAATAATTAACCAAGAAAAGCATCGTGTCAAT	1500						
Qy	1501	catttagaatttgataaagcgttaaatcgtgtcaagcgtattttaaccacaaacccacaaa	1560						
Db	1501	CATTTAGAATTTGATAAAGCGCTAAATGCTGTGTCAAGCGGTATTTAACCAAAACCCACAAA	1560						
Qy	1561	ctgaatttagggttggcttttgatcgatttaattcgtcttaggtcatcgtgggatatgact	1620						
Db	1561	CTGAATTTAGGGTTGGCTTTGATCGATTTAATTCGCTTATGGATCATGGGGATATGACT	1620						
Qy	1621	gcccaataataccaaaggcgttataccacgtaccgcggttagagggcgttttagataatcca	1680						
Db	1621	GCCCAATATACCAAGCGGTTATACGAGCTACCGCGGTAGAGGGCGTTTTAGATAATCCA	1680						
Qy	1681	tattttatcgcgcgcatccacgcagtatgaaacggtatcttttgttaataatacacgc	1740						
Db	1681	TATATTTATCGCCGCGATCCACGCAGTATTGAAACGGTATCTTTGTGTAAATAACACGC	1740						
Qy	1741	ggcgacatcttaaaactgtgaaccgcgtaaaaattaaagcgtagccattttgttagcttc	1800						
Db	1741	GCGCACATCTTAAACTGTGAACCGCGTAAATTAAGCGGATAGCCATTTTGTAGCTTC	1800						
Qy	1801	cgcgacttagtgaagcgagtagtggtattgggattaggggtcgcttttgatacaaat	1860						
Db	1801	CGCGATCTAGTGATAAGCGAGTATGTGGATTGGGATTAGGGGTGCGTTTTGTATCAACAT	1860						
Qy	1861	cgatttaaatcgtatgatccgtggacacttagccgaacttatcgaaattgctcttggaa	1920						
Db	1861	CGATTTAAATCTGATGATCCGTGGACACTTAGCCGAACCTTATCGAAATTTGCTTTGGAAT	1920						
Qy	1921	gggtgggattacgcttaaaaccaacagagtttgtatcgtcttcttctatcgcatctcaaacgt	1980						
Db	1921	GGTGGGATTACGCTTAACCAACAGAGTTTGTATCGCTTTCTTATCGCATTTCAACCGGT	1980						
Qy	1981	tttagagtcgctgattcttatgaactttatggttaaacgtgatcatatttgggcttaagat	2040						
Db	1981	TTTAGAGTGCCTGCATTTCTATGAACCTTATGTTAAACGTGATCATATTGGGCTTAAAGAT	2040						
Qy	2041	aacgaataatgtgaacgcgcgcaacgtagccacagtttagagccagaaaaatcgcactaat	2100						
Db	2041	AACGAATAATGTGCAACGCGCAACGATAGCCACCGATTAGAGCCGAGAAAAATCGACTAAT	2100						
Qy	2101	catgagattggagtagcttttaaagggtcaatttgggttaccttgatgtgagctattccgt	2160						

```
Db 2101 CATGAGATTGGATTAGCTTTAAAGGTCATTTGGTTACCTTGA-----ATTTCGGT 2152
Qy 2161 aataactataaaatatgattcgacagcatgtataaagaataatacaaaaatcacactgt 2220
Db 2153 AATAACTATAAAATATGATTCGCACAGCATGTAAAGAAATAATACAAAATCACACTGT 2212
Qy 2221 ttctataactaccataaattatcaagatgtacacactaaacgggataaaattttagtcgctaaa 2280
Db 2213 TTCTATAACTACCAATTAATTAAGATGTAGCACTAAACGGGATAAAATTTAGTCGCTAAA 2272
Qy 2281 ttgacttacacgggtattttatctatctgctccagatggtttttattcatcagttgcttat 2340
Db 2273 TTTGACTTACACGGTATTTATCTATGCTGCAGATGCTTTTATTCATCAGTTGCTTTAT 2332
Qy 2341 aacggtgtaaaagtaaaagacggaacactaacccgactcaagactcgatagcgtataacgat 2400
Db 2333 AACCGTGTAAAGTAAAGAGCGGAACCTAACCGACTCAAGACTCGATAGCGTAAACGAT 2392
Qy 2401 cctattctagatgcgattcagccagcagcgtatgtctgtgattgctggtacgatacccca 2460
Db 2393 CCTATTCTAGATGCGATTTCAGCCACGCGCTATGTCTTGGATTGCGCTACGATCACCCA 2452
Qy 2461 gaagaaaaatgggaattggcattactaccacattctaaagccaaaacgcccagatgag 2520
Db 2453 GAAGAAAAATGGGAATTGGCATTACTACCACTTATCTAAAGCCAAAACGCCGATGAG 2512
Qy 2521 gtggcagcagcacgtcatcagcgatatacgcgttgattagtggtggcaaacctgaccggtt 2580
Db 2513 GTGGCAGCAGCACGTCATCAGCGATACATCGGTTGATTAGGTGGCAAACTGACCGGTT 2572
Qy 2581 cttggtcacccatgatattaccggtttacatcaattataaaaactacaccccttaacgtggag 2640
Db 2573 CTTGGTACACCATGATATTACCGGTTACATCAATTAFAAAAACCTACACCTTACGTGGAG 2632
Qy 2641 gaattataatgtactaatcgtaaatattccacttgggaatcagtcgcccactcccggtg 2700
Db 2633 GAATTTATAATGTGACTTAATCGTAATAATTCACCTTGGGAATCAAGTCGCGCAATCCGGTG 2692
Qy 2701 tgaatgagtaaaacaaagacggggtagcaattacactcgatttggcgctccggggagaa 2760
Db 2693 TGAATGAGTAAACCAAGACCGGGGTAGCAATTACATCTCGATTGCGCTCCGGGGAGAA 2752
Qy 2761 atttcagtttagcatttgaaatgaagtttttag 2792
Db 2753 ATTTTCAGTTTAGCATTTGAAATGAAGTTTAG 2784

RESULT 2
T67241
ID T67241 standard; DNA; 3023 BP.
AC T67241;
DE 08-SEP-1997 (first entry)
DT Pasteurella haemolytica transferrin binding protein tbpa gene.
KW Transferrin binding protein; tbpa; receptor; vaccine;
KW Pasteurellosis; ss.
OS Pasteurella haemolytica serotype A1 strain H196.
FH Key
FT 13 Location/Qualifiers
FT cds
FT /*tag= a
FT /codon_start= 13
FT misc_difference 1115
FT /*tag= b
FT /note= "base 1115 is given as s in the
FT specification"
PN W09720934-A1.
PD 12-JUN-1997.
PF 29-NOV-1996; CA0791.
PR 01-DEC-1995; US-008569.
PR 01-DEC-1995; CA-164274.
PA (LORV/) LO B Y C.
PA (POTV/) POTTER A A.
PA (SCHR/) SCHRYVERS A B.
PI LO RYC, Potter AA, Schryvers AB;
```

```
DR WPI; 97-319780/29.
PT Pasteurella haemolytica transferrin binding proteins - used in
PT vaccines for prophylaxis and treatment of infection caused by
PT Pasteurella spp.
PS Example 1; Fig 3; 130pp; English.
CC The tbpB (T67240) and tbpA (T67241) genes of Pasteurella
CC haemolytica serotype A1 strain H196 were sequenced from recombinant
CC clones obtd. from a P. haemolytica A1 gene library using a probe
CC produced by PCR amplification (see also T67237-39). The genes are
CC organised in an operon arrangement of tbpB-tbpa, the tbpB gene being
CC preceded by a putative promoter and regulatory sequences. The two
CC genes appear to be coordinately transcribed. Full-length coding
CC sequences (T67235 and T67236) have also been obtd. from amino acid
CC sequences for TbpA (W18061) and TbpB (W18062) have been deduced.
CC Sequence 3023 BP; 890 A; 616 C; 667 G; 849 T;
SQ

Query Match 48.0%; Score 1339.8; DB 1; Length 3023;
Best Local Similarity 91.0%; Pred. No. 0;
Matches 1610; Conservative 0; Mismatches 97; Indels 62; Gaps 15;

Qy 1 atgataatgaataatcatcatttcgctatttccacttgccttaacacagttatttgcct 60
Db 13 ATGATAATGAATAATCATCATTTTTCGCTATTTCACCTGTTGCCTTAACAGTGTATTGCT 72
Qy 61 cttctcatcatcacgtgctgcgactgaaataaaaaatcgaagaaaaataacgatcta 120
Db 73 CTTTCTCATTCATACGGTGTGCGACACGACGACGACGACGACGACGACGACGACGACG 132
Qy 121 gctgttctggtatgaagtattgtgacagagagccattatgctcacgacgacgacgaa 180
Db 133 GCTGTTCTGGATGAAGTTATTGTGACAGAGAGGCCATTATGCTCACGAACGTCANACGAA 192
Qy 181 gtaactggtctggggaagtagtgaaaaattatcacgaaatagtgaaaaatcaaatcttt 240
Db 193 GTAACCTGGCTTGGGGAAGTAGTGAATAATATATCACGAAATAGTAAATAATCAAAATCTT 252
Qy 241 ggtattcgtgatttaacctcgtatgacctgacctgatttcggtggtggtggaacaggtcg 300
Db 253 GGTATTCTGATTTAACTCGCTATGACCTGCTGATTTCGGTGTGGTGGTGGTGGTGGTGG 312
Qy 301 gcaagtagtgctatgcaattcgcaggtgtagataaaaacccgtgctcagcttacttctgtat 360
Db 313 GCAAGTAGTGGCTATGCCATTCCAGAGGTGATAGATAAAACCGTGTACGCTTACTTGTGAT 372
Qy 361 gggctacacaaagcgcacagttatcatacagtaggttcagatgctaatggtggtgcaatt 420
Db 373 GGGCTACACACGCGCACAGTTATCATACGCT-GGTTTCAGATGCTAATGGTGTGCAATT 431
Qy 421 aatgagattgattgaaacattcgttcattgattgattgattgattgattgattgattgatt 480
Db 432 AATGAGATTGAGTATGAAACATTCGTTCAATTGAGTTAAGCAAAAGGAGCAAGTCTTCGGC 491
Qy 481 gaattgctctggtgcgcagtggtggtgctgattggttctgctactaaagatgcgcaggtat 540
Db 492 GAATATGCTCTGGTGCSCATGGTGGTGTCTATTGTTTTCGTACTAAGATTCGCCACGAT 551
Qy 541 attattaagaggggcagcattgggggttagatagtagtaagacctcttattgcagcaaaaat 600
Db 552 ATTATTAAAGAGGGGCGACGATTGGGGCTTACATAGTAGAAGACCTCTTATGCCAGCAAAAT 611
Qy 601 agccattttttacagtctatcgacgctggtggtgagcggtggttttgaagcacttctt 660
Db 612 AGCCA-TTTTTCAGTCTATCGCAGCGGCTGGTTCAGCGGGTGGTGTTCAGACGACTTGT 670
Qy 661 attgcaactcacgcagcagcgttaagagacacaaattcattccgagcgaataataataaa 720
Db 671 ATTGCAACTCACGACACGCTAAGAGACCAAAATTCATTCCGAGGCAAAATCAATT---A 727
Qy 721 cataatattcggogtataaccggctgttgaaatcgctacgactttaccacaaatcccgac 780
Db 728 CATATTATTCGGCGTATACCGGGCTTTCAAAATCGCTACGACTTTAGCC-AATTCGCGAC 786
```


Db 546 AAAGTGCCTACAGTAGCAAAAATCAACAATGGTTAAATCTACTTGGCTTTTGGGGTACTC 605
Qy 638 cggggtgttgaaagcactgttattgaactcaccagcagcgttaaaagagacccaaaattc 697
Db 606 ACAATGGCTTTGATGCTCTTGTGATTTACATCACCGTGATGGTAAAGAAACGAAAGCTC 665
Qy 698 attccgagcgaataaataaataataattcggcgat-----aaaccggctttg 748
Db 666 ATAAAGATGCAGAGAGTGGTTCTCAGAAATATCACCCGAGTAGGAGTGGAAACCAACGAGC 725
Qy 749 aaatcgcgtacgactttaccacaaattccgcacagaatcctcctgagagatcctcctttaa 808
Db 726 TTGATACCTCAAAATAGATATATCGCGACGACGAATAATCAACATACCTTATGGCTGGTTTT 785
Qy 809 ttggtgaagatactgcccacattagattgacttactcctcggaagggttaagtgaacc 868
Db 786 TGATTAAGATGAATGTCACAGCTTAGATTTGACGCCAAACAGATGGCTAGGGTGACAA 845
Qy 869 gcgataattcccagtgagaacatttccggaatatacgcctgaagagcgcacacagcttg 928
Db 846 AAGATACGCCATCTTCCGTTCTTACCCTGAATATACTCTCGAGGAAAAACAGGCTTATG 905
Qy 929 agcagattccttatcgactgagcagctctcagcccagaatatacccggttaaaagatcgca 988
Db 906 AGAACCAAAACATATTACAGAGCGCTTAAATGCTCAGGATTTACACTGGTGAATATAGAG 965
Qy 989 ttgcacaaaaccccttagattacaagagtaattctgttttatgaagtttgctatcact 1048
Db 966 CTTTACCTGATCCGCTTAAATATAAATCTGATCTTGGCTGGTTAAATAGATACACAT 1025
Qy 1049 tcaactgctcattattcttggcgcaatcttagaagatacaaaaacacgcgtacgatccc 1108
Db 1026 TCTCTCCGAAACATATTATCGCTGCTACTTATGAACATATAGCAAAACAGGTTTACGACACC 1085
Qy 1109 gtgatatacaaacgccagcttac----- 1131
Db 1086 GAGATATACCTATATACCCGCTTATTGGCAACCATCGGATTTACTTAGAACTGGTAGAAAT 1145
Qy 1132 --tatacaaaagacgatattaacttatcacttagaactatgtttatgaagggaataa 1189
Db 1146 GGTATCCAATGAAATGCTTAAGGATTTATTCGTGATTAATGCTTTAGATGGTGTGCTA 1205
Qy 1190 ttttagatggcttagtgttcaagcgaagtcacctttaggttgcgtatagccatgtga 1249
Db 1206 TTGACTACTTTACGGAAGATGGTGAAATCATCAAAAGGTTTACGTTGGCAAAAGCTC 1265
Qy 1250 agttttttgatgaacgctcaccacaaacgctgttttaggattaccctataaaat---aaac 1306
Db 1266 GTTTTATTTGACGAGTGGGCACACTCGTGATCGCTTAGGTCGCTTTATATCGTTATACCAATC 1325
Qy 1307 cagagaataatcgcgtgttgatagcattaaactcagtcgcgataaacaagaattgaac 1366
Db 1326 AAGATGGAATTCGCTGATGATAGACTATCCTTGAGTTTCGATCAGCAAAAATTAAT 1385
Qy 1367 tatatagcggctacatcgctgtgattgtagcattatcctgtgtagataaaaattgcc 1426
Db 1386 TATCTACCCGCTTGAGAGAGAACAACCTCTCCGAAATATCCAACCATAGATGAATTTGCC 1445
Qy 1427 gcccgactttggataaaacttgggtctatgatatcgaactgagcgttaataattaccagaaga 1486
Db 1446 GTGCAACTCTTGTAAACTTTGGCTTCAACTTAAAAATGAGCAAAAGTTCTTTATGAAGAA 1505
Qy 1487 agcatcgtgctattcatttagaatttgaataagcgctaaatgctggcgaagcgctattta 1546
Db 1506 AACACGACACTATTACGCTCTCGTTAGATAAAACCGCTACAAACCGGATTTGG----- 1557
Qy 1547 accaaacccacaaactgaatttaggttggcttggatcgatttaactcgtcttaggac 1606
Db 1558 ----TAAACATCAATTAATATGTTATTAGGTTGAGACCGGTTTCAATTCACCTTAAAC 1613
Qy 1607 atgggggatgactgcccacataataccaaggcggttataccagctaccgcggttagaggc 1666

Db 1614 GCCAGAAATTTTGGTGAATTTTCTGTGGAACTTGGCATCTGTATCAGAGGTAAACGGTT 1673
Qy 1667 gtttagataatcccatatatttatcgcc-----gcgatccacgcagtatgaaacggtat 1720
Db 1674 ATAAAGATACACCTTACATCTATGAGCTTAAAGATCAGGCAATTTATAGTAAAAATGAAT 1733
Qy 1721 ctttgttaataatacacgcgcgcacatcttaaaactgtgaaccgcgtaaaaattaaagcgc 1780
Db 1734 GTGATTATAGTGGCACATTTCAGGTAGGGGCTGATTGTGCTACAAAGTAAATCAAGGGC 1793
Qy 1781 atagccattttgttagettccgcg---atcagtgataagcgcgtatgtgattgggat 1837
Db 1794 ATAATCACTACATCGCTCTGAGAGATAAATTTGCATACAAACCAAGTATTGTGATATTGTT 1853
Qy 1838 taggggtgcgttttgcatacaatcgatttaaaactgtagatccgtggacacttgcgcgaa 1897
Db 1854 TGGTTACCGTTTCGATAAGCATAAATTCGATAGCATCATCGCTGGGCAATCAAGGCG 1913
Qy 1898 cttatcgaaaattggtccttgggaatggtggattacgctttaaaccaacagagtttgcacgc 1957
Db 1914 ATTATAAAACAGTGCCTGGAATATTGGCATAGTCGCAAAACCAACGTCATTCCTATCGC 1973
Qy 1958 tttcttatcgcaatttcaaacgcgtttttagagtcgctgattcttatgaaactttatggtaaaac 2017
Db 1974 TCTCTTATCGAGCATCTCTGGCTTTAGAGTGCCAAGTTTCCAAGAGCTATTTGGCTTAC 2033
Qy 2018 gtgatcatattggccttaaaagataaagaaatgtgcaacgcgcgaacgtgagccacagt 2077
Db 2034 GTTATGATGGTGCATGAAGGCTCCAGCGATGCTTA-----CCAAAAACAGAGAAAT 2087
Qy 2078 tagagccagaaaaactactaaatcatgagattggagttgagtttagctttaaaggtcaatttggtt 2137
Db 2088 TATCTCTGAAAAANTCTTAAACCAAGAGGTTGCTGGACTTTCAAAAGTGATTTGGTG 2147
Qy 2138 acctgtatgtagctatttccgttaataactataaaaaatagattgcaacagcatgtaaaa 2197
Db 2148 TCGTTGAAGTCAGTATTATTCAAAAATGACTATAAGCAGTTAAATTGCTCCAGCAGA----- 2202
Qy 2198 gaataatacaaaaaactcacactgtttctataactaccataataattcaagatgtagcaata 2257
Db 2203 -AAGATGCTCAACATCAATCAATGATTAACATATTATTTAATGTGCAAGATATTAAATTGG 2261
Qy 2258 acgggataaaattagtcgctaaatttgcattcacgcgttattttatctatgctgcagatg 2317
Db 2262 ACGCATTAATCTTATGGTAAGCTAGATTGGATGGGTTATTGATAAAATTCCTGAGG 2321
Qy 2318 gttttatcctcagttgcttataaacgcgtgtaaaagtaaaagcgcgaaataaccgcact 2377
Db 2322 GCATTTACACAACATTTGGCTTATAGCAAAATGCGAGTAAAGAGGTGAA---AAACTATC 2378
Qy 2378 caagactcgtatagcgttaaacgactccttctagatgcgattcagccagcagcgtatgac 2437
Db 2379 AAGGTTATGAAATTCGTTCTCCATTTGTAGATACCATTCAGCCTGCTCGCTATGTTG 2438
Qy 2438 ttggtatcggctacgctacccagaaagaaatggggaattggcattactaccacctatt 2497
Db 2439 TAGGAGTGGGTACGATCAGCCACATGAATAATGGGCGTGAATCTAACATGACACACT 2498
Qy 2498 ctaaaagccaaaaacgcgcgatgaggtggcagggcacacgtcatcgcgatacatcgcgttga 2557
Db 2499 CCAGTGGAAAAAATCCAAATGAGTTAAGAGGTAATGAACAGTCCGTTTGGCAATATG 2558
Qy 2558 tttagggtggc-----aaactgaccggttcttggtaacccccatgattaccgggttaca 2610
Db 2559 AGCGAACTGCCACGAAGAAAAAGAACTTTCTTGGCATACCTTTGACTTAAACGGGATATA 2618
Qy 2611 tca---attataaaactcacaccttgcgtgaggaatttataatgtgactaatcgttaaa 2667
Db 2619 TCACCCCTTGGAAACATACAAACGTACGAGTGGCTGCGTATATAACTGATGAATATCGTT 2678
Qy 2668 attccacttgggaatcagtcgcgaatccgggtgtgaatgcagtaaaccaaacgcgg---g 2724
Db 2679 ACACCACTTGGGAATCCGTACGTCAATCTTCGCTTAATGCAATTCATCAGCATACTAACG 2738


```
Db 3386 ATATTACGAAATAAGAACAAAGCGGCATCATTTGACAAAGCAGTGTAAAGTCTTAATC 3445
Qy 1352 aacaagatatgaactatatagccggctacatcgcttgcattgttagcgattatctctg 1411
Db 3446 AACAAAACATCATCTTGACAGTTATATGCAACATACGCAATGCGAGTCTTTATCCTAATC 3505
Qy 1412 tagataaaattgcgcgcgactttggataaattggtctatgtatgcgaactagcgta 1471
Db 3506 CAAGTAAGAAATGCGGCCCAACAGCTGATAAACCTTATTCATATATCATCTCTGATAGAA 3565
Qy 1472 ataattccaagaagaagcgtgctcattcatttagaatttgataaagcgcgtaaatgctg 1531
Db 3566 ATGTTTATAAGAAAAACATAATATGTTCGAATTGAATTTAGAGAAAAAAATTTCAACAAA 3625
Qy 1532 gtcaaggcgtatttaaccacaccccaaaactgaatttagg-----ttgggctttg 1582
Db 3626 ATTGGCTTACTCATCAAAATGCTCTCAATCTTGTTTGTGAGTACTTTACTTCAGCGCTTC 3685
Qy 1583 atcgatttaattcgttatgatatcggtgggatagctgcccataataccaaaagcggtt 1642
Db 3686 AGCATAAAGATTATTTAACTCGACGTGTACCGTACGGCAAGAGTATTTTCAGAGAAAG 3745
Qy 1643 ataccagctaccgggtagagggggttagataaaccatatttatcgcgcgcatccac 1702
Db 3746 CTAATGAACAAAGAGAAATGGTTACAAAAAACACCTTACTTATACCCAAAACCAACAG 3805
Qy 1703 gcagttatgaacgggtatctttgtgtaataacacg-----cgggacatcttaaac 1756
Db 3806 TAGGTTTGTAGTACAAGATCATTTGTGATTATTAAGGTAACTCTCTAAATFACAGAGACT 3865
Qy 1757 gtgaacgcgttaaaattaaagcgatagcgaattttgttagcttcgc---gatctagtga 1813
Db 3866 GTAAAGTGCGGTTAATTAAGGGAAAAATTAATTTATTCGACGACGCAATTAATATGGCAT 3925
Qy 1814 taacgagatgtgattggtgattaggggtgcgttttgatcaacatcgatttaaacctg 1873
Db 3926 TAGGAAATAGCTTGATTTAGTTTAGTTAGTATCGGTATGAGCTATCTCGCACAAAAGCTA 3985
Qy 1874 atgatacgtggacacttagccgaacttatcgaaattggtcttggaaatggtgggattacgc 1933
Db 3986 ATGAATCAACTATTAGTGTGTTGTAATTTAAATAATTTCTCTGGAATACTGGTATTGTCA 4045
Qy 1934 ttaaccacacagagttgtatcgctttcttctatcgattcattcaaaaggttttagagtgctg 1993
Db 4046 TAAACCAACAAAGGAATGGCTTGATCTTCTTATCGCCTTTCTACTGGATTAGAAATCCTA 4105
Qy 1994 cattctatgaacttttgtaaacgtgatcatattggcgttaagataaagaataatgtgc 2053
Db 4106 GTTTTGTGAAATGTATGGTTGGCGGTATGGTGCAATAATA----- 4147
Qy 2054 aacgcgcgaacgttagccaccagtttagagccagaaaaatcgactaatcatgagattggag 2113
Db 4148 ---CGAGGTTTATGTAGGTAATTTAAGCCTGAAACATCTCGTAACCAAGAGTTTGGTC 4204
Qy 2114 ttagctttaaaggccaatttggttaccttggatgtagctatttccgtaataactataaaa 2173
Db 4205 TCGCTCTAAAAGGGGATTTTGGTAATATTGAGATCAGTCATCTTTTAGTAATGCTTATCGAA 4264
Qy 2174 atatgttgagacgatgtaaa---agataatacaaaaaatacaactgtttctataact 2230
Db 4265 ATCTATCGCTTGTCTGAGAACTTAATAAAAATGGAAGCTGGAAGGCCAATTTAGAT 4324
Qy 2231 accataattcaagatgtagcactaaacgggataaaatttagtcgtcaaaatttgacttac 2290
Db 4325 ATCATATGCACAAAATGCAAAATTTAGTTGGCGTAATATAAATGCGCAATTAGATTTA 4384
Qy 2291 acggtatttatctatcgtccagatgggtttttattcatcagttgcttataacogtgtaa 2350
Db 4385 ATGGTTTATGGAACGATTTCCCTTACCGTGGTATGCAACATTTGCTTATAACCGAGTAA 4444
Qy 2351 aagtataaagcggaataactaacgactcaagactcgatagcgtataaagatccctatttag 2410
Db 4445 AAGTTAAGATC-----AAAAATCAATGCTGGTTTGGCCTCGGTAAAGCAGTTATTTATTG 4501
```

```
Qy 2411 atcgattcagccagcagcgtctatgtcttggattcggctacgatacccccagaagaaaaa 2470
Db 4502 ATGCCATTAGCCCGCCAGCGTTTATATCATTTAGTGGTTAGGCTATGATCATCCAAGTAATCTT 4561
Qy 2471 ggggaattggcattactaccacacctattctaaagccaaaacgcgcgatgagtgccagga 2530
Db 4562 GGGGAATTAATACAAATGTTTACTCAATCAAAAGCAAAATCTCAAAATGAATTGTAGGAA 4621
Qy 2531 caegtcatcagcggtacatcg-----gttgatttaggtggcaaacctgacctgttctt 2583
Db 4622 AACGTGATTTGGGTAAACAATTCAGGGATGTAAATCAACAGAAAATTTACTCGGGCAT 4681
Qy 2584 ggtacacccatgatattaccgg---ttacatcaattataaaaactacaccttacctgtgag 2640
Db 4682 GGCATATCTTAGATGTATCGGGTTTATACATGGCGAATAAAATATTATGCTTCGATTAG 4741
Qy 2641 gaatttataatgtgactaatcgtaaatattccacttgggaatcagtcgcgaatcccggtg 2700
Db 4742 GGATATATAATTTTAACTATCGCTATGTTACTTGGGAAGCGGTGCTGCTCAACAGCAC 4801
Qy 2701 tgaatgcagtaaaaccaaagacgg---ggtagcaattacactcgatttggcgctccggga 2757
Db 4802 AAGTGGCGGTCAATCAACATCAAAATGTTGGTAGCTATATCTGCTACGCGATCAGGAC 4861
Qy 2758 gaaatttcagtttagcatttgaatgaagttaa 2791
Db 4862 GAAACTATACCTTAACATTAGAAATGAATCTTA 4895
RESULT 5
V21431
ID V21431 standard; cDNA; 5099 BP.
AC V21431;
DT 20-JUL-1998 (first entry)
DE H. influenzae strain PAK 12085 transferrin receptor operon.
KW tbp1; tbp2; vaccine; H. influenzae; antibody; diagnosis;
KW passive immunisation; ds.
KW Haemophilus influenzae.
FH Key Location/Qualifiers
FT CDS 160..2124
FT /*tag= a
FT /product= Tbp1
FT 2152..4893
FT /*tag= b
FT /product= Tbp2
PN US5708149-A.
PD 13-JAN-1998; 487890.
PF 07-JUN-1995;
PR 08-NOV-1994; US-337483.
PR 08-NOV-1993; US-148968.
PR 29-DEC-1993; US-175116.
PR 07-JUN-1995; US-487890.
PA (CONN-) CONNAUGHT LAB LTD.
PI Chong P, Gray-Owen S, Harkness R, Klein M, Loosmore S,
PI Murdin A, Schryvers A, Yang Y;
DR WPI; 98-100410/09.
DR P-PSDB; W53048, W53049.
PT Purification of recombinant Haemophilus transferrin-binding protein
PS - by solubilising inclusion bodies separated from cell lysate
PS Example 4; Fig 6; 261pp; English.
CC The H. influenzae transferrin receptor operon contains two genes (tbp1
CC and tbp2) under the transcriptional regulation of one promoter. The
CC proteins encoded from these genes can be expressed in a recombinant
CC host. The proteins can be used in vaccines against H. influenzae
CC infections or to produce antibodies for use in diagnosis or passive
CC immunisation.
SQ Sequence 5099 BP; 1834 A; 849 C; 960 G; 1456 T;
Query Match 12.2%; Score 340.8; DB 1; Length 5099;
Best Local Similarity 50.4%; Pred. No. 8.3e-88;
Matches 1347; Conservative 0; Mismatches 1152; Indels 175; Gaps 14;
```

Qy	173	aaacgaagtaactggcttggggaaagtagtgaataattatcacgaataatgataaaatc	232
Db	234	ATAATGAAGTAACGTGACATTGGCAAAATATCAAAACAGTAGTGAAGTATCAGCCGAGAAC	2401
Qy	233	aaattcttgattcgttgatttaactcgtatgacctcgttatctcggttggtggaacaag	292
Db	2402	AGGTATTAATAATTTCTGTGATCTAACACGCTATGATCCAGGCATTTCAGTTGTAGAACAAG	2461
Qy	293	gtcgcggtgcaagtagtggctatgcattcaggtgtagataaaacacgctgtcagcttac	352
Db	2462	GCCGTGTCGAAGTTCTGGATATCTATTCGTGTATGACAGAAATAGAGTTGCTTTAT	2521
Qy	353	ttgttgatgggtaccacaagcgacagttatcatacgtctagggttc-----	398
Db	2522	TAGTAGATGGTTTACCTCAAGCGCAATCTTATGTAGTGCAAAGCCCTTTAGTTGCTCGTT	2581
Qy	399	----agatgctaagtgtgtgcaataataatgagtttgatgataaaacattcgttcaattg	454
Db	2582	CAGGATATCTGCGACTGTGTCAATTAATGAATTAATGAATATGAANAATGAAGCCGCTCG	2641
Qy	455	agttgaagcaaaagagcaagtcttcgcgaatatggctctgtgtgcacatgggtggtctattg	514
Db	2642	AAATAAGCAAGGGGGAGTTCTCTGAGTATGTTAATGAGCACTAGCTGGTCTCTGTAA	2701
Qy	515	gttttcgtactaaagatgcgaggatattataaagagggcgacatgggggttagata	574
Db	2702	CATTTCAAGCAAAATCAGCAGCGGATATCTTAGAAGGAGACAAATCATGGGGAATTCAAA	2761
Qy	575	gtaagacctcttatgcagcaaaaatagccatttttacagctctatcgacgcgctgggtg	634
Db	2762	CTAAAAATGCTTATTCACGCAAAATAAAGGCTTTACCATCTTTAGCTGTAGCTGGAA	2821
Qy	635	agcgcggtgtgtttgaagcactgttatgtcaactcacccgacacggttaaagagagcaaaa	694
Db	2822	AACAAGGGGATTTGACGGGGTCGCATTTTACTCAACGAAATTCATTTGAAACCCAAAG	2881
Qy	695	ttcttccagcgcaataaattaaaacataatattcggcgtatcaacggctttgaaatc	754
Db	2882	TCCATAAAGATGCATTAAAGGCGTACAAAGTTATCATCGATTAAATCGCC-----AAACC	2936
Qy	755	gtcacgactttaccaaaattccgcacagaatgctcctggaggatcctcttttaattgttg	814
Db	2937	AGAGGATCAATCTGCATCTTTGTGTGATGCAAGATGAGTGTCCAAGCCAGATGATTATA	2996
Qy	815	aagatactgcccacaattagattgactcctcgtgcgaagggttaagttgaaccgcgata	874
Db	2997	CAGTGTGTACCTTTC-----	3012
Qy	875	atttcccagtgagaaatttccggaaataacgctgaagagcgcaaacacgcttgagcaga	934
Db	3013	-----GCCAAGACCTTCGATTT	3031
Qy	935	ttccttatcgactgagcgctctcagcccagaataataccgggtaaagatacgtgaattgac	994
Db	3032	TATTCCTCCCAAGAGAAACCGTAAGCGTTTTCAGATTATACGGGGCTTAACCGTATCAAC	3091
Qy	995	caaaccttagattacagagtaattctgtttttatgaagtttggcgtatcaactcaact	1054
Db	3092	CTAATCCAATGAANAATGAAGCCAGTCTTGGTTTTTAAGAGGAGGTATCATTTTTCGTG	3151
Qy	1055	cgtctcattatcttggcgcaattcttagaagatacaaaaacacgctacgatccgtgata	1114
Db	3152	AACAACATTAATTTGGTGGTATTTTGAATTCACACAAACAAAAATTTGATATCGTGATA	3211
Qy	1115	tgcacaagccagcttactatacaaaagacgatattaaactataccttaggaactatgttt	1174
Db	3212	TGACATTTCCCGCTTATTTAAGATCAACAGAAAAACGGGATGATAGCAGTGGCTTCTTTT	3271
Qy	1175	atgaaggggataataatttagatggcttagtgttcaagcccaaggatcccttattgggttgc	1234
Db	3272	ATCCAAAGCAAGATTATGTTGATATCAACGATTGAGGATGGCCG-----AGGCGTTA	3325

Qy	1235	gctatagccatgtgaagtttttttgatgaacgtcaccacaaacgtcgtttttaggtattcaacct	1294
Db	3326	ACTATGCAAGTGGCGTTTATTTCGATGAACACCATAGAAAAACGCGTGTAGGTATTGAAT	3385
Qy	1295	ataaataaaacacagagaataatcgcctgggttgatagattaaa---ctcagtcgaggata	1351
Db	3386	ATATTTACGAAATATAGAACAAAGCGGCATCATTTGACAAAGCAGAGTGTTAAGTGTCTAATC	3445
Qy	1352	aacagatatgaactatataagccggtcacatcgcttcgattgtagcgtattatctcttgg	1411
Db	3446	ACACAAACATCATACTTGACAGCTATATATGCAACATACGCATTGCGACTCTTTATCCTTAATC	3505
Qy	1412	tagataaaaaattccgcccgcgaactttggataaaatcttggctctatgatcgaactgagcgtga	1471
Db	3506	CAAGTAAGAATTCCGCCCCACACAGCTGATAAACCTTATTTCATCTATCATCTTCGTATGAA	3565
Qy	1472	ataataccaagaaaaagcctgctcattcattttagaatttgataaagcgcgtaaatgctg	1531
Db	3566	ATGTTTATNAAGAAACAACTAATATGTTGCAATTGAAATTTAGAGAAAAAATTCACACAA	3625
Qy	1532	gtcaaggcgtatttaaccacaccccaactgaaatttaggg-----ttggggcttg	1582
Db	3626	ATTGGCTTACTCATCAAAATGCTTCATCTCTGGTTTGTGATGACTTTACTTCACGCGTTC	3685
Qy	1583	atcgatttaattcgtcttagatcgatcggggatagctgcccataataccaaaggcggtt	1642
Db	3686	AGCATAAAGATTATTTAACTCGACGCTGTACCGCTACGCCAAAGAGTATTTACAGAGAAAG	3745
Qy	1643	ataccagctaccggttagaggcgcttagataaatacctatatatttatcgcgcgatccac	1702
Db	3746	CTAATGAAACCAAGAAGAAATGGTTACAAAAACAACTTACTTATACCCAAAAACCAACAG	3805
Qy	1703	gcagtatgaaacggtatcttctgtataataacacg-----cgggcgacatcttaact	1756
Db	3806	TAGTGTTTGTGATACAGAGTCAATGTGTGATTTAAAGGTAACCTCTCTAATTTACAGAGACT	3865
Qy	1757	gtgaacgcgcgtaaaaataaaggcgaagcatttggtagcttcgcg---gatctagtga	1813
Db	3866	GTAAGTGCGGTTTAATTAAGSGMAAAATTAATTATTTCCGACGACGCAATAATATGCGAT	3925
Qy	1814	taecgagatgtggaatttgggattaggggctgcttttgatacaacatcgatttaaatctg	1873
Db	3926	TAGGAAATACGTTGATTTAGTGTTAGTATTCGGTATGACGTATCTCGCACAAAGACTA	3985
Qy	1874	atgataccgtggcacacttagccaaactatcgaaattgcttgaatgggtggattacgc	1933
Db	3986	ATGAATCAACTATTAGTGTGGTAAATTTAAAAAATTTCTCTTGGAACTACTGGTATTGTCA	4045
Qy	1934	ttaaacacacagagtttgtatcgcttctctatcgcatctcaaacgggttttagatgcctg	1993
Db	4046	TAAACCAACGGAATGCGTGTATCTTCTTATCGCCTTTCCTACTGGATTATAGAAATCCTA	4105
Qy	1994	cattctatgaacttttatgttaaacgtgatacatatcttgggctttaagagataacgatctgc	2053
Db	4106	GTTTGTGTGAATGTATGTTGGCGGTATGGTGGCGGTATGGTGGCAATAATA-----	4147
Qy	2054	aacgcgcgcgaactagaccacagtagaccagagaaaaatcgactaaatcatgagattggag	2113
Db	4148	---CGGAGGTTTATGATAGTAAATTTAAAGCCCTGAACATCTCGTAAACCAAGAGTTGGTC	4204
Qy	2114	ttagctttaaaggtcaatttggttaccttgatgtgagctatttccgtaataactataaaa	2173
Db	4205	TCGCTCTAAAAGGGGATTTTGGTAATATTGAGATCAGTCAATTTAGTAACTCTTATCGAA	4264
Qy	2174	atatgattgcagacatgtaaa---agaataatacaaaaatcnaactgtttctataact	2230
Db	4265	ATCTATTACGCTTTTGCTGAAGAACTTAATATAAAATGGAACCTGGAAGGCCAATATTGGAT	4324
Qy	2231	accataattcaagatgtagacctaacaacgggataaatttagtcgctaaatttgacattac	2290
Db	4325	ATCATATGCAAAATGCAAAATTAAGTTGGCGTAAATATAACTGCGCAATTAGATTTTA	4384
Qy	2291	acggtattttatctatgctgccegatgggtttttatctcatcagttgctttaaaccggttaa	2350

```

Db 4385 ATGGTTTATGGAACGATATTCCTACGGTGTGATGCAACATTTGCTTATAACCGAGTAA 4444
Qy 2351 aagtaaaagagcggaaactcaacgactcaagactcgatagcgtaaacgactctattctag 2410
Db 4445 AAGTTAAAGATC--AAAATCATATGCTGGTTGGCCCTCGGTAGCAGTATTTATTGG 4501
Qy 2411 atgcgattcagccagcagcgtatgcttgattcggttagcagatcacccagaagaaaaa 2470
Db 4502 ATGCCATTACGCCAGCCAGCGTTATATCATTTGTTAGGCTATGATCATCAAGTAACTT 4561
Qy 2471 ggggaattggcattactaccacatttctaaagcaaaacgcgcgagatgaggtggcaggca 2530
Db 4562 GGGGAATTAATACAAATGTTACTCAATCAAAAGCAAAATCTCAAAATGAATGCTAGGAA 4621
Qy 2531 cagctcatcagcgatcacatcg-----gttgattgagtggtgcaaacctgaccggtctt 2583
Db 4622 AACGTGCATTGGGTACAAATTCAGGGGATGTAATAATCAACAGAAATTTACTCGGGCAT 4681
Qy 2584 ggtacacccatgatattaccgg---ttacatcaattataaaactacaccttacctggag 2640
Db 4682 GGCATATCTAGATGATATCGGTTATACATGGCGAATAAAAAATATTATGCTCGATTAG 4741
Qy 2641 gaattataatgtactaatcgtaaatattccacttgggaatcagtgcgccaatccggtg 2700
Db 4742 GGATATATAATTTATWCAACTATCGCTATGTTACTTGGGAAGCGGTGCGTCAAAACAGCAC 4801
Qy 2701 tgaatgcagtaaaacaaagaccgg---ggtagcaattacactcgatttggcgtccggga 2757
Db 4802 AAGGTGCGGTCAATCAACATCAAAATGTTGGTAGCTATATCTCGGTACGACATCAGGAC 4861
Qy 2758 gaaatttcagtttagcattggaatgaagttaa 2791
Db 4862 GAAACTATACCTTAACATTAGAAATGAATCTA 4895

```

RESULT 6

```

ID Q94445
AC Q94445;
DT 09-OCT-1996 (first entry)
DE Bacterial transferrin receptor operon (H. influenzae strain PAK 12085).
KW Tbp1; Tbp2; transferrin receptor operon; vaccine; antigen;
KW non-typable strain; Haemophilus influenzae; meningitis; ss.
OS Haemophilus influenzae strain PAK12085.
FH Location/Qualifiers
FT cds 160..2124
FT FT /*tag= a
FT FT /product= Tbp2
FT FT 2152..4896
FT FT /*tag= b
FT FT /product= Tbp1
FT FT 4891..4893
FT FT /*tag= c
FT FT /note= "no amino acid residue given in
FT FT corresponding protein"
PN W09513370-A1.
PD 18-MAY-1995.
PF 07-NOV-1994; CA0616.
PR 08-NOV-1993; US-148968.
PR 29-DEC-1993; US-173116.
PR (CONN-) CONNAUGHT LAB LTD.
PI Chong P, Gray-owen S, Harkness R, Klein M, Loosmore S;
PI Murdin A, Schryvers A, Yang Y;
DR WPI: 95-194089/25.
DR P-PSDB: R77889-90.
DT Nucleic acids encoding Haemophilus transferrin receptor - used to
PT develop prods for detection and in diagnosis, prevention and
PT treatment of Haemophilus infection.
PS Claim 12; Fig 6A-Q; 231pp; English.
CC The present DNA shows the transferrin receptor (Tfr) operon consisting of
CC two genes (Tbp1 and Tbp2) arranged in tandem and which are transcribed
CC from a single promoter, from the non-typable Haemophilus influenzae

```

```

CC strain PAK 12085. H. influenzae Tfr is iron- and/or haemin-regulated and
CC a putative fur-binding site has been identified upstream of tbp2.
CC Antibodies blocking this binding site may prevent bacterial growth.
CC Fragments of the Tfr (or its genes) are useful in vaccines to provide
CC protection against, e.g. bacterial meningitis. An advantage of using the
CC Tfr is that it shares homology with Tfr of other H. influenzae strains
CC including non-typable strains.
SQ Sequence 5099 BP; 1834 A; 854 C; 955 G; 1456 T;

```

```

Query Match 12.1%; Score 339.2; DB 1; Length 5099;
Best Local Similarity 50.3%; Pred. No. 2.4e-87;
Matches 1346; Conservative 0; Mismatches 1153; Indels 175; Gaps 14;

```

```

Qy 173 aaaaagaaagtaactgcttgggaaagtagtaaaattatcacgaatgagtaaaatc 232
Db 2342 ATAATGAAGTANCTGGCAACTTATCAAAACTAGTGAAGATATACGCCGAGAAC 2401
Qy 233 aaattcttggtattcgtgatttaactcgctatgacctggtatttcggtggtgaaacaag 292
Db 2402 AAGTATTAAATATTCGTGATCTAACACGCTATGATCCAGGCATTTTCAGTTAGACAAAG 2461
Qy 293 gtcggtgcaagtagtggctatccattcgaggtgtagataaaaaacggtgcagcttac 352
Db 2462 GCCGTGTGCAAGTCTGGGATATTCATTCGTGTATGGACAGAAATAGAGTTGCTTTAT 2521
Qy 353 ttgttgatggtgctaccacaagcgcacagtattatcaccgttaggttc----- 398
Db 2522 TAGTAGATGGTTTACCTCAACGCAATCTTATGTAGTGAAGGCCCTTTAGTTGCTCGTT 2581
Qy 399 ----agatgctcaatggtggtgcaattaatgattgagttgagtaaaacattcgttcaattg 454
Db 2582 CAGGATATTCGGCACCTGGTCAATTAATGAATTTGAATGAATGAATGAATGAATGAATGA 2641
Qy 455 agttaagcaaaagagcaagttcttcggaatgctgctggtggtggtggtggtggtggtggtg 514
Db 2642 AAATAAGCAAGGGGGGAGTTTCTTCTGAGTGTGTAATATGAGCAGTACTAGTGGTCTGTAA 2701
Qy 515 gtttcgtactaaagatgctgcagcagattatttaaagaggggcagcattggggttagata 574
Db 2702 CATTTCAAGCAATATCAGCAGCCGATATCTTAGAAGGAGACAAATCATGCGGAATTCAAA 2761
Qy 575 gtaagacctcttgcagcaaaatgacctttttlacagctctatcgagcggctggtg 634
Db 2762 CTAAAAATGCTTATTAAGCAAAATTAAGCCTTTACCCATCTCTTAGCTGTACTGGAA 2821
Qy 535 aggcgggtggttttgagcactgtttattgcaactcaccgacacggttaagagagaccaaaa 694
Db 2822 AACAGCGGGATTTGACGGGTGCGCCATTTATATCAACGAAATTCATTTGAACCCCAAG 2881
Qy 695 ttcatccgaggaataataataataatatttcggcgtataacccggtttgaaatc 754
Db 2882 TCCATAAAGATGCTTAAACGGCTACAAAGTTATCATTCGATTAAATCGCC-----AAACC 2936
Qy 755 gctacgactttaccataattccgcacagaatgctcctggaggatctccttttaattggtg 814
Db 2937 AGAGGATCAATCTGCATATCTTGTGATGCAAGATGAGTGTCCAAAGCCAGATGATTATA 2996
Qy 815 aagatacttgcacacacattagattgtactcctcgtgcaaggttttaagttaaacccgcgata 874
Db 2997 CAGTTGTTTACCTTTC----- 3012
Qy 875 atttccagtgagaaacatttccgggaatatagcctgaagagcgcaaacagcttgagcaga 934
Db 3013 -----GCCAAACGACCTCGGATTT 3031
Qy 935 ttcttatcgactgagcagctctcagcccaagaataatccggttaagatcgcttgcac 994
Db 3032 TATCTCTCCCAAGAGAAACCGTAAAGCTTTTCAGATTATACGGGGCTTAACCGTATCAAC 3091
Qy 995 caaaccttttagattacaagagtaattctgtttttatgaagtttggctataccttcaact 1054
Db 3092 CTATCCAATGAATATGAAAGCCAGTCTTGGTGTGTTTAAAGAGGAGGATATCATTTTTCTG 3151

```

Qy	1055	cgctcattatcttggcgcaattcttagaagatacaaaaacacgctaagatatccgtgata	1114
Db	3152	AACAACATTATTGGTGGTATTTTGAATTTCACACAACAAAAATTGATATCGTGTATA	3211
Qy	1115	tgcgaacgccagcttactatatacaaaagacgatataaactctacttaggaactatgttt	1174
Db	3212	TGACATTTCGCCGTATTAAAGATCAACAGAAAAACGGGATGATAGCAGTGGCTCTTTT	3271
Qy	1175	atgaaggggataatatttagatggccttagtgttcaagccaaggatcccttaggggtgc	1234
Db	3272	ATCCAAAGCAAGAT-----TAGGTGCATATCAAGCTATTGAGGATGCCGAGGCGTTA	3325
Qy	1235	gctatagccatgtgaagtttttttgatgaacgtcaaccacaaacgcttgtttaggattcaact	1294
Db	3326	ACTATGCAAGTGGCTTTATTTTCGATGAACACCATAGAAAAACAGCGTGTAGGTATGCAAT	3385
Qy	1295	ataaataaaaccagagaataa---tcgcgtgttgtagcattaaactcaagtgcggata	1351
Db	3386	ATATTTACGAAAAATAAGAACAAAGCGGCATCATTTGACAAAGCAGTGTTAAGTCTAATC	3445
Qy	1352	aaceagatatgaactatatagccggctacatcgcttgcttagcgttagcgattatcctctgg	1411
Db	3446	ARCAAAACATCATACTGTGACAGTTATATGCAACATACCGCATTTGCAGTCTTTATCTCAATC	3505
Qy	1412	tagataaaattgcgcgcgcactttggataaaactcttgctctatgatacgaaactgagcgta	1471
Db	3506	CAAGTAGAATTCCGCCCCCAACACGCTGATAAACCTTATTCATCTATCATCTCTGATAGAA	3565
Qy	1472	ataattaccaagaagacgtgctgctatctcatattagaatttgataaagcgtcaaatgctg	1531
Db	3566	ATGTTTATAAGAAAAACATAATATGTTGCAATTGAAATTTAGAGAAAAAAATTCACAAA	3625
Qy	1532	gtcaagcgctatttaaccaaaaccccaaacactgaatttaagg-----ttgggctttg	1582
Db	3626	ATTGGCTTACTCATCAAAATGTCTTCAATCTTGGTTTGTGATGACTTTACTTCAGCGCTC	3685
Qy	1583	atcgatttaactgccttatggatacgtgggatacgactgcccataataccaaagcggtt	1642
Db	3686	AGCATAAAGATTATTAACTCGACGTGTACCGCTACGGCAAGAGTATTCAGAGAAAG	3745
Qy	1643	ataccagctaccgggtagaggcgcttagataatccatatattatcgccgcgagctccac	1702
Db	3746	CTAATGAAAAAAGAAGAAATGTTACAAAAAACAACTTACTTATACCCAAAACCAACAG	3805
Qy	1703	gcgctattgaacggctatcttgtgtaataatacacg-----cgggacactcttaaac	1756
Db	3806	TAGTTTTGTGTAGTACAAGATCATGTGTGATTTAAAGGTAACCTCTCTAATTACAGAGACT	3865
Qy	1757	gtgaaccgcgtaaattaaaggcgatagccatttggtagcttcgcg---gattctatga	1813
Db	3866	GTAAGTGCCTGTTAATTAAAGGMAAAATATTATTTCGACAGCAGCAATAATATGCGAT	3925
Qy	1814	taagcggagtatggtatttgggattagggggtgcgttttgatacaacatcgatttaaactg	1873
Db	3926	TAGGAAATACGTTGATTAGTTTAGGTTTTCGGTATGCGTATCTCGCACAAAAGCTA	3985
Qy	1874	atgatcgtggacacittagccgaacttatcgaaattggtcttggaatgggtgggattacgc	1933
Db	3986	ATGAATCAACTATTAGTGTGTGTAAATTTAAAAAATTTCTTGTGAATACGTGATGTCA	4045
Qy	1934	ttaaaccaacagagttgtlatcgcttctctatcgcatctcaaacggttttagagtgctg	1993
Db	4046	TAAACCAACCGAATGGCTTGATCTTCTATCGCTTCTACTCTGATTAGAAATCCTA	4105
Qy	1994	cattctatgaactttatggtaaacgtgatcatatttgggcttaaagataaagaaatagtgc	2053
Db	4106	GTTTTGCTGAATGTATGTTGCGGTATGTTGGCAATAATA-----	4147
Qy	2054	aagcgcgcgaactagcccaccagtttagagcgcgaaaaaatcgactaaatcatgagattggag	2113
Db	4148	---GCGAGGTTTATGTAGGTAAATTTAAGCCCTGAAACATCTCGTAAACCAAGAGTTTGGTC	4204

Qy	2114	ttagctttaaaggtcaatttggttacccttgatgtagagctatttccgcaataactataaaaa	2173
Db	4205	TCGCTCTAAAAGGGGATTTTTGGTAATATTAGATCAGTCACTTTTAGTAATGCTTATCGAA	4264
Qy	2174	atatgattggacagcatgtaaa---agaataatacaaaaatacacactggtttctataact	2230
Db	4265	ATCTATTTCGCTTTTGCTGAAGAACTTAATAAAATGGAACCTGGAAAGGCCCAATATTATGGAT	4324
Qy	2231	accataataatcaagatgtagcactaaaacgggataaaatttagtcgcgtataaatttgacttac	2290
Db	4325	ATCATATATGCACAAATATGCAAAATATAGTTGGCGTAAATATAAATTCGCCCAATTAGATTTTA	4384
Qy	2291	acggtattttatctatctgctgccagatgggttttttattcatcctagttgcttataaacctgttaa	2350
Db	4385	ATGGTTTTATGAAACCGTATTCCCTACGGTTGGTATGCAACATTTGCTTATAACCGGACTAA	4444
Qy	2351	aagtaaaagcgggaactaaacgcgactcaagaactcgatagcgtataaacttacttctag	2410
Db	4445	AAGTTAAAGATC---AAAAAATCAATGCTGGTGTGGCTCCGTAAAGCAGTTATTATTATTG	4501
Qy	2411	atcgacttcagccagcacgctatgcttgcttgcttcgctacgataccccagaagaaaaat	2470
Db	4502	ATGCCATTACGCCACCGGTTATCATTTGGTTTGGCTATGATCATCCCAAGTAATACTT	4561
Qy	2471	ggggaattggcattactacacacttctaaagccaaaacgcgcgatgaggtggcaggca	2530
Db	4562	GGGGAATTAATACAATGTTTACTCAATCAAAAGCAAAATCTCAAAATCAATTGCTAGGAA	4621
Qy	2531	caogtcatcgcgatacatgc-----gtgatttgagggtggcaaaactgacoggtctt	2583
Db	4622	AACGTGCATTTGGGTAAACAATTCAAAGGGATGTAAAAATCAACAAGAAAAACTTACTCGGCAT	4681
Qy	2584	ggtacacccatgatattaccgg---ttacatcaattataaaactacaccttaoigtgag	2640
Db	4682	GGCATATCTTAGATGTATCGGGTTATTTACATGGCGAAATAAAATATTATGCTTCGATTAG	4741
Qy	2641	gaatttataatgtgactaatcgtataatctccacttgggaatcagtcgccaatccgggtg	2700
Db	4742	GGATATATAATTTATTCAACTATCGCTATGTTTACTTTGGGAAGCGGTGGTCAAAACACAC	4801
Qy	2701	tgaatgcagtaaaccaagaccgg---ggtagcaattacactcattgttggcgtccgggga	2757
Db	4802	AAGGTGGCGGTCAATCAACATCAACATCAAAATGTTGGTAGCTACTCTCGTACGACATCAGGAC	4861
Qy	2758	gaaatttcagtttagcatttgaatgaagtttta	2791
Db	4862	GAAACTATACCTTAACATTAGAATGAATTTCTA	4895
RESULT	7		
TA9504			
ID	TA9504	standard; DNA; 5144 BP.	
AC	TA9504;		
DT	05-MAY-1998	(first entry)	
DE	Transferrin receptor gene sequence encoding Tbp1 and Tbp2.		
KW	Transferrin receptor; Haemophilus influenzae type b;		
KW	iron; human transferrin; iron source; antibody; bacterial growth;		
KW	vaccine; immunogenic truncated analogue; antigen; Tbp1; Tbp2; ss.		
OS	Haemophilus influenzae.		
FT	Key	Location/Qualifiers	
FT	CDS	192..698	
FT	FT	/*tag= a	
FT	FT	/note= "encodes Tbp2 (W08966)"	
FT	FT	2135..4870	
FT	FT	/*tag= b	
FT	FT	/note= "encodes Tbp1 (W08965)"	
PN	W09640929-A2.		
PD	19-DEC-1996.		
PF	07-JUN-1996; CA0399.		
PR	17-MAY-1996; US-649518.		
PR	07-JUN-1995; US-483577.		
PA	(CONN-) CONNAUGHT LAB LTD.		
PI	Chong P, Gray-Owen S, Harkness RE, Klein MH, Loosmore SM,		

Qy	1814	taagcagagtatg	tggaattt	gggattag	gggg	tgctt	gtgatacaaca	tcgatt	taaatctg	1873
Db	3897	TAGGAAATACAT	TGATTTAGGTTT	TAGGTATTCGGTAT	TTCGGTATGAGCTAT	CTCTACGACAAAAGCTA	3956			
Qy	1874	atgacgctgg	acacctatag	ccgaacttac	gaaattg	ctctt	gaaatgg	gggaattacgc	1933	
Db	3957	ATGANTCAACT	ATTAGTGTGTGTA	AAATTTTAA	AAATTTCTCTT	TGGAACTACGTTATGTCA	4016			
Qy	1934	ttaaccaacagag	tttgtatcgctt	ctcttcatcg	ctatactcaaac	cggttttagag	tgctg	1993		
Db	4017	TAAACCAACG	GAATGGCTTGATCTT	CTTCTATCGCCTT	CTACTGGATTTAGAAATCCTA	4076				
Qy	1994	cattctatga	actttatggt	taaacgtgat	catatctgg	gctttaagagata	caagaatatgtgc	2053		
Db	4077	GTTTTGCTG	AAATGATATGTT	TGGCGGTATG	TGGCAATAATAG	CGGATTTATGTAGGTA	4136			
Qy	2054	aacgcgcga	actagtagcc	ccaggttag	agccagaa	aaatcgact	aaatcatgagat	tgga	2113	
Db	4137	AA-----	TTTAGCCTG	AAACATCTCGT	TAAACCAAGAGTTTGGTC	4175				
Qy	2114	ttagcttta	aagtc	aaatttg	ttacctgtat	gtgagctat	ttccgtta	aaactataaaa	2173	
Db	4176	TCGCTCTAAA	AAGGGGATTTTGGT	TAATTTAGAT	TATGAGTCAGTCA	TTTAGTAATGCTTATCGAA	4235			
Qy	2174	atatga----	ctgcgacag	catgataa	aagaataac	aaaaatcac	actgtttctata	2227		
Db	4236	ATCTATTAC	CGCTTTCGTG	GAAGAACTTTAGT	TAAAAATGGAAC	TACTGGAAAGGGCAATTTATG	4295			
Qy	2228	actaccata	atatcaaga	tgtag	cactaaac	gggataaa	tttagtcgct	laaaatttgact	2287	
Db	4296	GATATCAT	ATATGCAC	AAATGCAAAAT	TAGTTGGCGT	AAATATAACT	TCCGCAATTAGATT	4355		
Qy	2288	tacacgg	gctatttat	ctatctat	gtctgcga	tggtttttat	ctcatcag	tgtcttataacc	2347	
Db	4356	TTAATGCTT	TATGGAAC	CGTATTCCCT	CTACGGTTGTTATG	CAACATTTGCTTTATAACCGGAG	4415			
Qy	2348	taaaagt	taaaag	agcagga	actaac	ccgactcaag	actcgatag	cgtaaaacgactctattc	2407	
Db	4416	TAAGA	TTAAGATC----	AAAAATCA	ATGCTGTGTTGGCCT	CCGTAGACAGTTATTTAT	4472			
Qy	2408	tagatgc	gatttc	agccag	cagcgtat	gtcttggatt	ctcggtac	gatcacccagaga	2467	
Db	4473	TTGATGCC	ATTACGCC	CAGCGTTATATCAT	TGTTTAGGCTATGATCAT	CCAAGTAATA	4532			
Qy	2468	aatggg	gaattg	gcattact	caccactat	ctaaaggcc	aaaaacgcg	atgaggtggcag	2527	
Db	4533	CTTGGG	GAATTAATAC	AATGTTTACT	TCATCAAAAGCAAAATCT	CAAAATGAATTTGCTAG	4592			
Qy	2528	gcacagc	atcatc	gatacat	gcg-----	gttgattag	gtgg	caaacctgaccggtt	2580	
Db	4593	GACAACT	GTGATTTGGGT	TAACATTC	AAGGAATGTAA	NAATCAACA	AGAAACTTTACT	CGGG	4652	
Qy	2591	cttgg	tacaccc	atgatatt	accgg---	ttacat	caattata	aaaactacac	ctttag	2637
Db	4653	CATGGC	ATATCTTAGAT	GTATCGGTTATTAC	ATGCGAAATA	AAAAATATTATG	CTTCGAT	4712		
Qy	2638	gaggaa	tttataat	gtgact	aaatcg	taaaatat	ctcact	ttgggaat	cagtcg	2697
Db	4713	TAGGGAT	TATAATTTATT	TATCAACTAT	TCGGTATTTACT	TTGGGAAG	CGGTGGCT	CAAAACAG	4772	
Qy	2698	gtgtga	atcag	taaac	ccagacc-----	gggtag	caattac	actctga	atttgg	2754
Db	4773	CACAAG	TGGGTC	CAATCAAC	ATCAAAATGTTGGT	AGTATATCT	CGCTAC	GCACAT	CAG	4832
Qy	2755	ggagaa	atttcag	tttag	cat	ttgaa	atga	agt	tttta	2791
Db	4833	GACGAA	ACTATAC	CTTAAC	ATTAGA	ATGAATGA	ATTCTTA	4869		

RESULT	8
V21440	
ID	V21440
AC	V21440

DT	20-JUL-1998	(first entry)
DE	H. influenzae strain SB33 transferrin receptor operon.	
KW	tbp1; tbp2; vaccine; H. influenzae; antibody; diagnosis;	
KS	passive immunisation; ds.	
OS	Haemophilus influenzae	
FH	Key Location/Qualifiers	
FT	CDS 192..698	
FT	FT /tag= a	
FT	FT /product= Tbp1	
FT	FT 2135..4870	
FT	FT /tag= b	
FT	FT /product= Tbp2	
PN	US5708149-A.	
PD	13-JAN-1998.	
PF	07-JUN-1995; 487890.	
PR	08-NOV-1994; US-337483.	
PR	08-NOV-1993; US-148968.	
PR	29-DEC-1993; US-175116.	
PR	07-JUN-1995; US-487890.	
PA	(CONN-) CONNAUGHT LAB LTD.	
PI	Chong P, Gray-Owen S, Harkness R, Klein M, Loosmore S,	
PI	Murdin A, Schryvers A, Yang Y;	
DR	WPI: 98-100410/09.	
DR	P-PSDB: W54123, W54124.	
DR	Purification of recombinant Haemophilus transferrin-binding protein	
PT	- by solubilising inclusion bodies separated from cell lysate	
PS	Example 4; Fig 7; 261pp; English.	
CC	The H. influenzae transferrin receptor operon contains two genes (tbp1	
CC	and tbp2) under the transcriptional regulation of one promoter. The	
CC	proteins encoded from these genes can be expressed in a recombinant	
CC	host. The proteins can be used in vaccines against H. influenzae	
CC	infections or to produce antibodies for use in diagnosis or passive	
CC	immunisation.	
QY	Sequence 5144 BP; 1813 A; 888 C; 974 G; 1469 T;	
QY	Query Match 11.9%; Score 332.6; DB 1; Length 5144;	
DB	Best Local Similarity 50.5%; Pred. No. 1.9e-85;	
DB	Matches 1353; Conservative 0; Mismatches 1134; Indels 190; Gaps 15;	
QY	173 aaacggaagtaactggcttggggaaagtagtgaaaaattatcacgaaatgagataaaatc 232	
DB		
DB	2325 ATAATGAAGTAACGTGGACTTGGGCAAAATATATAAAACGAGTGAAGATATCAGCCGAGAAC 2384	
QY	233 aaattcttggtatcgtgatttaactgcgtatgccctggatttcggtgggtgaaacaag 292	
DB		
DB	2385 AAGTATTAATATTCGTGATCTACACGCTATGATCAGGCATTTTCAGTGTGAACACAG 2444	
QY	293 gtccggtgcgaagtagtggcttatgccattcggaggttagataaaaaacggtgcagcttac 352	
DB		
DB	2445 GTCGGGTGCAAGTTCGGATATTCATTCTGCTGCTATGGACAGAAATAGAGTTGCTTTAT 2504	
QY	353 ttgttgatgggtaccacaaagcgacagttatcatcagctagggttc----- 398	
DB		
DB	2505 TAGATAGTGGTTTACCTCAACGCAATCTTATGATGTCNAAGGCCCTTTAGTTCGCTT 2564	
QY	399 ---agatgctaaagggtgcgaattaatgagattgagatgataaaacatttcgccaattg 454	
DB		
DB	2565 CAGGATATTCGGCAGTGGTGCATTAATGAATGTAATGAAATGTAAAGGCCGTCG 2624	
QY	455 agtLaagcaagggagcaagttctcggaatatggctctggtgcgcatggtggtgctattg 514	
DB		
DB	2625 AAATAAGCAAGGGGGGNGTCTTCCTGNGATGATGTAATGGAGCACTAGCTGCTGTAA 2684	
QY	515 gtttcgtactaaagatgcgcaggtattattaaagaggggcagcattgggggcttagata 574	
DB		
DB	2685 CATTTCAAGCAAAATCCGACGCCGATCTCTTGAAGAGGACAAAATCATGCGGGAATTCAAA 2744	
QY	575 gtaagacctcttatgcagcaaaaaaatagccattttttacagtctatgcagcagcgtggtg 634	
DB		
DB	2745 CTAAAAATGCTTTATTCAGGCAAAAATAAAGGCTTTTACCATTCTTTAGCTGTAGCAGGAA 2804	
QY	635 agcggggtgggttttgaaagcactttatttgcaactcaccgcacggtataaaagaccgaaa 694	

Db 2805 ACAAAGTGGATTGAAGGGGTCGCCATTTACACTCAACGAATTCGGAGGAACCCAAAG 2864
QY 695 ttctccaggagcaataaataaataaataattcgcggtatataacccggtttgaaatc 754
Db 2865 TCCATAAAGATGCATTAAGAAGCGTACAAGTATGAGCGATTCATCGCC-----ACAACA 2920
QY 755 gtaacgactttaccacaaatccgcacagaatgctcctcgtgagatcctcttaattgtgg 814
Db 2921 GATAAATCTTCAGGATACATTTGTGATACAAGGTGAGTG-----TCCAAATGGTG 2969
QY 815 aagatacttgcccaacattagattgtactcctcgtgcaagggttaagttgaaccgcgata 874
Db 2970 ATGACAAGTGTGCAGC----- 2985
QY 875 atttccagtgagacatttccggaataatacgcctggaagcgcgaacagcgttgagcaga 934
Db 2986 -----CAAAACCCTGCAAGT 3002
QY 935 ttcttatcgactgagcagctctcagcccaagaataatacoggttaagatcgcatgcaac 994
Db 3003 TATCCCCCAAGCGAAACCGTAAAGCGTTTACAGATTATACGGGGCTAACCGTATCAAAAC 3062
QY 995 caaaccttttagattacaagagtaattcgttttttatgaagtttggtctatacactcaact 1054
Db 3063 CTAATCCCAATGAATATGAAGCCAGCTGTGGTTTTTAAGAGAGGGGTATCATTTTCTG 3122
QY 1055 cgtctcattatcttgccgaactcttagaagatacacaacacgcgtacacatccgtgata 1114
Db 3123 ACAACACTATATGGTGTATTTTGAATTCACACACAAAAATTTGATATCCGTGATA 3182
QY 1115 tgcaacgcgcagcttactatacaaaagcagatattaacttatcacttaggaactatgttt 1174
Db 3183 TGACATTTCCCGCTTATTTAAGATCAACAGAAAAACGGGATGATAGAACTGSCCCTTTT 3242
QY 1175 atgaaggggataaatttttagatgcttagtgttcaagcccaaggatcccttatgggtgc 1234
Db 3243 ATCCAAAGCAAGATTATGGTGTATATACAGTATTTGAGGATGGCCGA-----GGCGTGA 3296
QY 1235 gctatagccatggaagttttttagtaagcgtcaccacaacgcgtttaggattcaact 1294
Db 3297 ACTATGCAAGTGGGCTTATTTTCGATGAACACCATAGAAAACAGCGTGTAGGTATTGAAT 3356
QY 1295 ataaataaacaacagagaataatcgtcgtgttgatagcattaaactc---agtcgggata 1351
Db 3357 ATATTTACGAAAAATAAGAAACAAAGCGGCATCATTTGACAAAACAGCTGTTAAGTGTAAATC 3416
QY 1352 aacaagataatgaactatatacggcgctacatcgcgttgccattgtgagcattatcctgtg 1411
Db 3417 AAAAAACATCATCTTGACATTTATATGCGACATATGCGACATACGCATTTGCTTTATCCTAATC 3476
QY 1412 tagataaaaaatgcccgcgactttggataaaatcttgcctatgtagaactgagcgata 1471
Db 3477 CAAAGTAAGATTTCCCGCCGACACTGATGAACCTTTATCATTATCTGTTCTGATAGAA 3536
QY 1472 ataatacaagaagaacatcgtgtcattcatttagaatttgataaagcgtctaaatgctg 1531
Db 3537 ATGTTTATAAGAAAAACATAATAATGTTGCAATTGAATTTAGAGAAAAAAATTCACAAA 3596
QY 1532 gtaaggggtatttaaccaaaccacaactgaaatttggg-----ttgggctttg 1582
Db 3597 ATGTGCTTACTCATCAATTTGCTCAATCTTGGTTTTGATGACATTTACTTCAGCGCTTC 3656
QY 1583 atcgatttaactcgtttatggtatcgtgggataatagctgccaataatacacaaggcgtt 1642
Db 3657 ACATAAAGATTTATTAACCTGACGTGTTACCGCTACGGCAAAATATTTATTCAGGGAACAG 3716
QY 1643 ataccagctaccggttagagggcgttttagataatccatatatttatcgccgcgcatccac 1702
Db 3717 TTGCTGGTAAACGAAGAAATGTTACGAAAAACACCTTACTTTACTCAAAACCAAAAG 3776
QY 1703 gcagttgaaacggttatcttttgtgtaataatacacgcggc-----gacatctttaaact 1756

Db 3777 TAGATTTGTAGACAAGATCATTTGTAATATATAAAGTAGTCTCTTAATTACACGACT 3836
QY 1757 gtgaaccgcgttaaaattaaagcgatagccatttggtagcttcgcg---gatctagta 1813
Db 3837 GTAAAGTCGGTTAATTAAGGGAAAAATTAATTTCGCAGCAGCAATAATATGGCAT 3896
QY 1814 taagcgagtagtggatttgggatttaggggtgcgttttggatcaacatcgatttaactcg 1873
Db 3897 TAGGAAAAATCATTTAGTTAGTTTAGTATTCGGTATGACGTATCTCGTACAAAAGCTA 3956
QY 1874 atgacctggagacacttagcgaacttatcgaaatttggcttcttggaaatgggtggaatacgc 1933
Db 3957 ATGAATCAACTATTAGTCTTAAATTTAAAAATTTCTCTGGAAATCTGCTATTGTCA 4016
QY 1934 ttaaaccaacagagtttgcgtcttctatcgcttctcaacggttttagagcgctg 1993
Db 4017 TAAAAACCAACGGAATGGCTTGTATCTTTATCGCTTTTCTAGTGAATTTAGAAATCCTA 4076
QY 1984 cattctatgaactttatgtaacggtgatcatattgggcttaaaagataaacgaatatgtgc 2053
Db 4077 GTTTTGTGTAATGATGTTGGGGTATGTGGCAATAATAGCGATCTTTATGTAGTA 4136
QY 2054 aacgcgcgaacgtagccaccagtttagagccagaaaaatcgaactaatcatgattgag 2113
Db 4137 AA-----TTTAAAGCCTGAAACATCTCGTAAACCAAGAGTTTGGTC 4175
QY 2114 ttagctttaaaggtcaatttggttacccttgatgagctatttcogtaataactataaaa 2173
Db 4176 TCGCTCTAAAAGGGGATTTGGTATATTTAGATCAGTCATTTTAGTAAATGCTTTATCGAA 4235
QY 2174 atatga-----ttgcgacagcatgtaaaagaataatacaaaaaatcacactgtttctata 2227
Db 4236 ATCTTATCGCCTTTGCTGAAGACTTAGTAAATGGAACACTACTGGAAGGCAATTAAG 4295
QY 2228 actacataattcaagatgtagcaactaaacgggataaaatttagtgcataaattgact 2287
Db 4296 GATATCAATAATGCACAAAAATAGTTGGCTAAATATAACTGCGCAATTAGATT 4355
QY 2288 taccggtattttatctatgctgcagatggtttttatctatcagttgcttataaccgtg 2347
Db 4356 TTAATGGTTTATGGAACGTTATCCCTACGGTTGGTATGCAACATTTGCTTTATAACCGAG 4415
QY 2348 taaagtataaagagcggaaacttaaccgactcaagactcgatagcgtaaacgactctattc 2407
Db 4416 TAAAGTTAAAGATC---AAAAATCAATGCTGGTTGCGCTCCGTAGCAGTATTATTAT 4472
QY 2408 tagatgcattcagccagcagctatgctgtggtatcggtacgatacccccagaagaaa 2467
Db 4473 TTGATGCCATTTCAGCCACGCGTTATATCATTTGTTTATAGCTATGATCATCAAGTAATA 4532
QY 2468 aatggggaatggcattactacccttattcctaaagccaaacccgcgataagggtggcag 2527
Db 4533 CTGCGGGAATTAATACAATGTTTACTCAATCAAAAGCAAAATCTCAAAATGAATTTGCTAG 4592
QY 2528 gcaacgctcatcaggtatcacatcgc-----gttgatttaggtggcaaacgactgacccgtt 2580
Db 4593 GACACGTCGATTCGGTAAACAATTCAGGAATTAATAATCAACAGAAACTTACTTCGGG 4652
QY 2581 cttggtaacccatgatatattaccgg---ttacatacaattataaaactacaccttactg 2637
Db 4653 CATGCAATCTTATAGATGATCGGTTTATACATGGCGAATAAAAAATATTATGCTTCGAT 4712
QY 2638 gagaaattataaagtgaactaatcgtaaatccacttgggaatcaatggcgcaactcg 2697
Db 4713 TAGGATATAAATTAATCAACTATCGCTATGTTACTTTGGGAAGCGGTGCGTCAAAACAG 4772
QY 2698 gtgtgaatgcagtaaaaccaagacc---ggggtagcaaatcacactgatttggcgctcgg 2754
Db 4773 CACAGGTGCGGTCAATCAACATCAAAATGTTGGTAGCTATACTCTGCTACGACGATCAG 4832
QY 2755 ggagaaatttcagtttagcatttgaatgaagttaa 2791
Db 4833 GACGAACTATACCTTAACATTAGAAATGAATTTCTA 4869

Db 3459 CAAGTAAGAATTGCGGCCCAACACATTGATAAACCTTATATCATATATCGTTCTGTATAGAA 3518
Qy 1472 ataattaccagaagaagcatcgtgcatcatttagaatttgataaa----- 1518
Db 3519 ATGTTTATAAGAAAACATATATGTTGCAATTCGAATTTAGAGAAAATAATTCACAAA 3578
Qy 1519 --gocgtaaatgctggtcaagcgttatttaaccacacccacacacgaatttagggttgg 1576
Db 3579 ATTGGCTTACTCATCAAAATGCTTCAATCTGTTGTTGATGACATTTACTTCAGCGGCTTC 3638
Qy 1577 gcttgatcgatttaattacgttattggtatcgatcggtggtatgactgcccataataccaaag 1636
Db 3639 AGCAATAAGATTATTAATACTGACGTGTTATGCTGACGCAATAGTATTCAGAGAAAC 3698
Qy 1637 ggggttataccagctaccgcggtagaggcgttttagataatccatataatttatgcgcgcg 1696
Db 3699 CTGCTGAAACTGGTAAACCAAGAATGTTTGCATATCAACCTTACTTATACCCAAAC 3758
Qy 1697 atccacgcagttattgaacgcgtattcttctgtgtataataacacgcggc-----gacatct 1750
Db 3759 CAGAGCCATATTTTGCAGGACAAGATCATTTGTAATATATCAAGGTAGCTCCCTTAATTACA 3818
Qy 1751 taaactgtgaacgcggtataaattaaagcgatagccatttggtagcttcgc---gac 1807
Db 3819 GAGACTGTAAAGTCGGTTAATTAAGGGAATAATTTATTTGCGACGACGCAATAATA 3878
Qy 1808 tagtataagcgagtatgtgatttggattaggggtcgttttgatcaacatcgattta 1867
Db 3879 TGGCATTAGGGAATACGTTGATTTAGTTTAGGTATTCGGTATGACGTATCTCGTACAA 3938
Qy 1868 aatcgtatgacgcgtggacacttagccgaacttatcgaaatttgcttgggaatggtggga 1927
Db 3939 AAGCTAATGAATCAACTATTAGTGTGTTGTAATAATTTAAAAATTTCTTGGAACTACTGTA 3998
Qy 1928 ttacgcttaaaccaacagatttgcgtcttcttcttcgcatttcaaacggttttagag 1987
Db 3999 TTGTCATAAAACCAAGGAATGGCTTGATCTTTCTTATCGCCTTTTCTACTGGATTAGAA 4058
Qy 1988 tgcctgcatcttatgaactttatgtaaacgctgatacatattgggtcttaagataacgaat 2047
Db 4059 ATCTAGTTTTCGAATATGATGTTTGGCGGTATGGTGGCAAGNATCACGAGGTTTATG 4118
Qy 2048 atgtgcaacgcgcgcaacgtagccaccagtttagagccgagaaaaatcgactaatcatgaga 2107
Db 4119 TAGGTAAA-----TTTAAGCCTGAAACATCTCGTAACCAAGAGT 4157
Qy 2108 ttggagttagctttaaaaggtcaattgtgtacccttgatgtagctatttcctcgtaataact 2167
Db 4158 TTGGTCTCGCTCTAAAGGGGATTTTGGTAATATTGAGATCAGTCAATTTTAGTATGCTT 4217
Qy 2168 ataaaaatgatgtgcacagcatgtaa---agaataatacaaaaatcacactgttct 2224
Db 4218 ATCGAAATCTTATCGCCTTTCGTAAGAACCTTAGTAAAAATGGAATCGGAAGGGCAATT 4277
Qy 2225 ataactaccataataatccaagtagtagcactaaacgggataaaaattagctcgtaaaattg 2284
Db 4278 ATGGATATCATAAATGCACAAAATGCAAAATAGTTTGGCGTAAATATAACTGCACAAATAG 4337
Qy 2285 acttacaagggtatttatactatgctgcagatggtttttattcatcatcagttgcttataacc 2344
Db 4338 ATTTTAAATGGTTTATGGAAACGATATTCCTACGGTGTGTTGTAACACATTTGCTTATAACC 4397
Qy 2345 gtgtaaaaagtaaaagcggaactaaacgcactcaagctgataagcgaatcgaatccta 2404
Db 4398 AAGTAAAAGTTAAAGATC---AAAAAATCAATGCTGGTTTAGCCTCCCTGATAGCAGTTAT 4454
Qy 2405 ttctagatgcgattcagccagcacgctatgtgcttgggtattcgttcacatcacccagaag 2464
Db 4455 TATTTGATGCCATTACGCCCGCGTTATATCATTTGTTTGGTTATGATCATATCCAGATA 4514
Qy 2465 aaaaatgggaattggcattactaccacattcttaagagccaaaacccgcatgaggtgg 2524
Db 4515 ATACTTGGGAATTAATCAATGTTTACTCAATCAAAAGCAAAATCTCAAAATGAATTGC 4574

Qy 2525 caggcacacgctcatcacggatacatcgc-----gttgatttaggtggcaaacgtgaccg 2577
Db 4575 TAGGAAAACGTGCATTAGGTAAACAATTCAAGGGGATGTAAATATCAACAAGAAATCTACTC 4634
Qy 2578 gttcttgtagacacccatgatattaccgg---ttacatcaattataaaaaacacacaccttac 2634
Db 4635 GGGCATGGCATATCTTAGATGTATCGGTTTATTACATGGCGAATAAAAAATATTATGCTTC 4694
Qy 2635 gtgaggaatttataatgtgactaatcgtaaatattccacttgggaatcagtgcgccaat 2694
Db 4695 GATTAGGGATATATAATTATTATCAACTATCGTATGTACTTGGGAACGGGTGCGTCAAA 4754
Qy 2695 ccggtgtgaatgcagtaaaaccaagaccg---gggtagaacaattcacactcgatttggcgctc 2751
Db 4755 CAGCACAGGTGCGGTCAATCAACATCAAAATGTTGGTAGCTATATCTCGTACGACGAT 4814
Qy 2752 cggggagaaaatttcagtttagcatttgaattgaatgaagtttta 2791
Db 4815 CAGGACGAAACTATACCTTAACATTAGAAATGAATTTCTA 4854

RESULT 10

V21430
ID V21430 standard; cDNA; 5009 BP.
AC V21430;
DT 20-JUL-1998 (first entry)
DE H. influenzae strain Minn A transferrin receptor operon.
KW tbp1; tbp2; vaccine; H. influenzae; antibody; diagnosis;
KW passive immunisation; ds.
FH Haemophilus influenzae.
FT Key Location/Qualifiers
FT CDS 121..2103
FT /tag= a
FT /product= Tbp1
FT 2117..4855
FT /*tag= b
FT /product= Tbp2
PN US5708149-A.
PD 13-JAN-1998.
PF 07-JUN-1995; 487890.
PR 08-NOV-1994; US-337483.
PR 08-NOV-1993; US-148968.
PR 29-DEC-1993; US-175116.
PR 07-JUN-1995; US-487890.
PR (CONN-) CONNAUGHT LAB LTD.
PI Chong P, Gray-Owen S, Harkness R, Klein M, Loosmore S,
PI Murdin A, Schryvers A, Yang Y;
DR WPI; 98-100410/09.
DR P-PSDB; W53046, W53847.
PT Purification of recombinant Haemophilus transferrin-binding protein
PT - by solubilising inclusion bodies separated from cell lysate
PS Example 4; Fig 5; 261pp; English.
CC The H. influenzae transferrin receptor operon contains two genes (tbp1
CC and tbp2) under the transcriptional regulation of one promoter. The
CC proteins encoded from these genes can be expressed in a recombinant
CC host. The proteins can be used in vaccines against H. influenzae
CC infections or to produce antibodies for use in diagnosis or passive
CC immunisation.
SQ Sequence 5009 BP; 1809 A; 836 C; 934 G; 1430 T;

Query Match 11.9%; Score 331.2; DB 1; Length 5009;
Best Local Similarity 50.5%; Pred. No. 4.7e-85;
Matches 1354; Conservative 0; Mismatches 1133; Indels 193; Gaps 15;

Qy 173 aaaaacgaagtaacttggttgggaaagtagtaaaaattatcacgaatgagtaaaaatc 232
Db 2307 ATATGAAGTAACCTGGACTGGCAAAATTTACAAACTAGTGAAGTATACGCGAGAAC 2366
Qy 233 aaattcttggtattcgtgatttaactgcgtatgacctggtatttcggtggtggaacaag 292
Db 2367 AAGTATTAAATATTCGTGATCTAACACGCTATGATGATCAGGGGATTTTCAAGTTGACACAAG 2426

Qy 293 gtcggtgcaagtagtggtgcctatgccattcgaggtgtagataaaaaacggtgcagcttac 352
Db 2427 GTCGGGTGCAAGTTCTGGATATTCTATTCGTGTATGGACAGAAATAGAGTTGCTTTAT 2486
Qy 353 ttgtgtgagggctaccacaacgacaggttatcatcagctaggctc----- 398
Db 2487 TAGTAGATGGTTTACCTCAACGCAATCTTATGTAGTGCAAAGCCCTTTAGTTCGTCGTT 2546
Qy 399 ----agatgctaattggtggtgcaattaatgagattgagatgataaaaaacattggtccaattg 454
Db 2547 CAGGATATTCGCACATGSGTCAATTAATGAATTAATGAATGAATAATGAATAAGCCGTCG 2606
Qy 455 agttaagcaaaagagcaagttctctcggaatatggtcctctggtgcgcacatggtggtgctattg 514
Db 2607 AAATAAGCAAGGGGGGAGTTCTCTGAGTATGTAATAGGAGCAGCTAGCTGTTCTGTAA 2666
Qy 515 gtttcgtactaaagatgctgcaggatattataaagaggggcagcattggggcttagata 574
Db 2667 CATTTCAAGCAATCAGCAGCCGATATCTTAGAAGGAGACAAATCATGCGGAATTCAAA 2726
Qy 575 gtaagacctctatgccaagcaaaaatagccattttttacagctctatcgcagcggctggtg 634
Db 2727 CTAAAAATGCTTATTCAAGCAAAAAATAAGGCTTTAGCCATCTTTAGCTGTAGCAGGAA 2786
Qy 635 agcgggtggttttgaagcacctgttattgcaactccagcacagcggtaaaagagacaaaa 694
Db 2787 AACAAAGTGGATTTGAAGGGCTAGCCATTTACACTCAACGAAATTCAAATGAAACCCCAAG 2846
Qy 695 ttcatcccgaggcaaaataataaaacataataatcggcgataacggctttgaaatc 754
Db 2847 TCCATAAAGATGCATTAAGCGGTACAAAGCTTATGATCGATTAAATCGCC----ACAACA 2902
Qy 755 gctaocgacttaccocaaatccgcacagaaatgcctcctggaggaatcctctttaaattggtg 814
Db 2903 GATAAATCTTCAGGATACTTTGTGATACAAGGTGAGTG-----TCCAAATGGTG 2951
Qy 815 aagatacttgcccaacattagatgtactcctcgtgcaagggttaagttgaacgcgcgata 874
Db 2952 ATGCAAGTGTGCAGC----- 2967
Qy 875 atttcccagtgagaacatttcgggaataatacgcctggaagagcgcaacagcttgagcaga 934
Db 2968 -----CAAGCCACTCGGACTT 2984
Qy 935 ttccctatgcgactgagcagctctcagccccagaataatcccggttaagatgcgattgcaac 994
Db 2985 TATCCACCACAAAGCGAAACCGTTAGCGTTTCAGATTTATACGGGGCTTAACGGTATCAAAAC 3044
Qy 995 caaaccttttagattacaagagtaattctgtttttatgaagtttggtctatcacctcaact 1054
Db 3045 CTATCCAAATGAATATGAAGCCAGTCTTGTTTAAAGAGGGGTATCATTTTTCGTG 3104
Qy 1055 cgtctcattatcttgccgcaattcttagaagatacaaaaacgcgtcacgatatccgtgata 1114
Db 3105 AACACATTTATTTGGTGTATTTTGAATTCACACACAAAAATTTGATATCCGTGATA 3164
Qy 1115 tgcnaacgccagcttactatacaaaagacgatattaaattatcacttaggaactatgttt 1174
Db 3165 TGACATTTCCGGCTTATTAAAGCCCAACAGAAAGCGGATGATAGTAGTGTGTTT 3224
Qy 1175 atgaagggaataatttttagatggcttagtgcagcgaaggtcccttatgggttgc 1234
Db 3225 ATCCAATGCAAGATCATGGTGATATCAACATATTTAGGAGTGGCA-----GAGGCGTTA 3278
Qy 1235 gctaagcactgtagtgaattttttagaagcgtccaccacaacgctgtttaggattcacct 1294
Db 3279 AATATGCAAGTGGGCTTATTTCGATGAACCATAGAAACACGCGTGTAGGTATTGAAT 3338
Qy 1295 ataaataaaacccagagaaataatcgtggttggatagcattaaatc---agtgcggata 1351
Db 3339 ATATTTAGCAAAATTAAGAACAAAGGGGGCATATTGACAAAGCAGTGTGAAGTGTCAATC 3398

Qy 1352 aacaagattgaactatatagcggctacatcgcttgcttagcgtattagcattatccctgtg 1411
Db 3399 AACAAAACATCATACTTTGACAGTTATATGCGACATAGCATTCAGTCTTTATCCTTAATC 3458
Qy 1412 tagataaaaattgcccgcgactttggaataaacttggtctatgtatcgaaactgagcgta 1471
Db 3459 CANGTAGAATTGCCGCCCAACACTTGATAAACCTTATTCTACTATCTGTTCTGATAGAA 3518
Qy 1472 ataattaccagaagaagcatcgctgctcatcatttagaatttgataaa----- 1518
Db 3519 ATGTTTATAAAGAAAAACATAATATGTTGCAATTTGAATTTAGAGAAAAAATTCACAAA 3578
Qy 1519 --gcgtaaatgctggttcaagggctatttaacccaaaccccaaaactgaatttaggttgg 1576
Db 3579 ATTGGCTTACTCATCAAAATGCTCTCAATCTTGCTTTTGATGACTTTTACTTCAGCGCTTC 3638
Qy 1577 gctttgatcgatttaattcgtcttatggtatcatgctgggataatgactgcccataataccaaag 1636
Db 3639 AGCATAAAGATTATTTAACTCGAGCTGTTATCGCTACGGCAGATAGTATTCCAAGGAAAC 3698
Qy 1637 gcggtataccagctaccgcggtagaggcggttttagataatcaatatttatcgccgcg 1696
Db 3699 CTGGTGAACCTGTTAAACCAAGAAATGTTTGCATCAACAACTTACTTTACCCAAAAC 3758
Qy 1697 atccacgcgatttgaacgggtatcttctgtgtaataacacacgcgcg-----gacatct 1750
Db 3759 CAGAGCCATATTTGCGAGGACAGATCATTTGTAATTAATCAAGGTAGTCTCTTAATTACA 3818
Qy 1751 taaacttgaaacgcgctaaataataaagcgatagccattttgttagcttcgcg---gatc 1807
Db 3819 GAGACTGTAAAGTCGGTTTAAATAAGGGAAAAATTTATTTCCGACGACGCAATAATA 3878
Qy 1808 tagtgataaagcagtagctggtattgggataaggggtaggggttttgatcaacatcgatga 1867
Db 3879 TGGCATTAGGGAATACGTTGATTTAGTATTAGGTATTCGGTATGACGTATCTCGTACAA 3938
Qy 1868 aatcgtatgcgctggacacttagccgaactatcgaaatgtgctctgaaagtggtgga 1927
Db 3939 AAGCTAATGAATCACTATTAGTGTGCTGAATTTAAATAATTTCTCTTGGAACTACTGGTA 3998
Qy 1928 ttacgcttaaaccaacagagttgtgtatcgcttcttctatcgcatcttcaaacgggttttagag 1987
Db 3999 TTGTCATAAAAACCAACGAATGGCTTGATCTTCTTATCGCTTTCTACTTGGATTTAGAA 4058
Qy 1988 tgcctgcattctatgaactttatggttaaacgtagatcatattgggcttaagataaacgaat 2047
Db 4059 ATCCTAGTCTTTCTGAAATGTATGTTGGCGGTATGTTGGCAAGAAATGACGAGGTTTATG 4118
Qy 2048 atgtgcaacgcgcgaacgtagccaccagttagaccagagaaataatcgactaatcatgaga 2107
Db 4119 TAGGTAAA-----TTTAAGCCTGAAACATCTCGTAACCAAGAGT 4157
Qy 2108 ttggagttagctttaaaggtcaatttggttacctgtatgtgagctatttccgttaataact 2167
Db 4158 TTGCTCTCGCTCTAAAAGGGGATTTTGGTAATATTGAGATCAGTCATTTTAGTAATGCTT 4217
Qy 2168 ataaaaatagattgacagcagcatgtaaa---agaaataacaaaaatcacactgtttct 2224
Db 4218 ATCGAAATCTTATCGCTCTTTCCTGAAGAACTTAGTAAAAATGGAACCTGGAAGGGCAATT 4277
Qy 2225 ataactaccataatattcaagatgtagcactaaacgggataaaatttagtgcgtcaaatgtg 2284
Db 4278 ATGGATATCATATGACAAAAATGCAAAATTAGTTGGGTAAATATTAAGTCACAAATTAG 4337
Qy 2285 acttacacgggtattttatctatgctgcagatggtttttattcatcagttgcttataacc 2344
Db 4338 ATTTTAATGGTTTATGGAACGTTATCCCTAGCGTTGGTATGCAACATTTTCTTTATAACC 4397
Qy 2345 gtgtaaaagtaaaagacgcggaactaacccgactcaagactcgatagcgtaaacatccta 2404
Db 4398 AAGTAAAAGTTAAAGATC---AAAAAATCAATGCTGTTTAGCCTCCGTACGAGAGTTATT 4454
Qy 2405 ttctagatgggattccagccagcagcgtatgctgtggattcgggtacgatacccccagaag 2464

Db 4455 TATTGTATGCCATTGAGCCAGCCGTTATATCATTTGGTTAGGCTATGATCATCAAGTA 4514
Qy 2465 aaaaatgggaattggcattactaccaccattctaaagccaaacacccgcatggtgg 2524
Db 4515 ATACTTGGGAATTAATCAATGTTTACTCAATCAAAAGCAAAATCTCAAAATGAATTGC 4574
Qy 2525 caggcacacgtcatcacgatacatcgc-----gttgattggtggcaaacgtgaccg 2577
Db 4575 TAGGAAAAACGTCATTAGTAGTAATCAAGGGATGTAATAATCAACAAAGAAACCTTACTC 4634
Qy 2578 gtctctgtatacccatgataattaccg---ttacatcaattataaaacatacacttac 2634
Db 4635 GGGCAATGCAATCTTAGATGATCGGGTTATACATGCGGAATAAAATATATATGCTTC 4694
Qy 2635 gtggagggaattataatgtgactaactgaaatattccacttgggaatcagtgccccaat 2694
Db 4695 GATTAGGATATATAATTAATCAACTATCGTATGTTACTTGGGAAGCGTGGCTCAA 4754
Qy 2695 ccggtgtgaatgcagtaaaccaagaccg---gggtagcaattatacactcgattggcgctc 2751
Db 4755 CAGCAACAAGGTGCGGTCAATCAACATCAAAATGTTGGTAGCTATATCTCGCTACGCAGCAT 4814
Qy 2752 cggggagaaattcagtttagcatttgaaatgaagtttta 2791
Db 4815 CAGGACGAACACTATACCTTAACATTAGAAATGAATTTCTA 4854

RESULT 11
T49501
ID T49501 standard; DNR; 5033 BP.
AC T49501;
DT 05-MAY-1998 (first entry)
DE Transferrin receptor gene sequence encoding Tbp1 and Tbp2.
KW Transferrin receptor; Haemophilus influenzae type b;
KW iron; human transferrin; iron source; antibody; Bacterial growth;
KW vaccine; immunogenic truncated analogue; antigen; Tbp1; Tbp2; ss.
OS Haemophilus influenzae.
FH Key Location/Qualifiers
FT CDS 169..2151
FT /*tag= a
FT /*note= "encodes Tbp2 (W08960)"
FT 2129..4903
FT /*tag= b
FT /*note= "encodes Tbp1 (W08959)"
W09640929-A2.
PD 19-DEC-1996.
PF 07-JUN-1996; CA0399.
PR 17-MAY-1996; US-649518.
PR 07-JUN-1995; US-483577.
PA (CONN-) CONNAUGHT LAB LTD.
PI Chong P, Gray-Owen S, Harkness RE, Klein MH, Loosmore SM,
PI Mordin AD, Schryvers AB, Yang Y;
DR WPI: 97-052329/05.
DR P-PSDB: W08959-60.
PT Haemophilus truncated transferrin receptor protein analogue, Tbp2
PT used to induce protection against disease caused by transferrin
PT producing pathogens, or as antigen to detect Haemophilus Tfr
PT antibodies
PS Claim 5; Fig 4A-Q; 228pp; English.
CC The present sequence represents the transferrin receptor gene of
CC Haemophilus influenzae type b, strain Pagan. The bacterial transferrin
CC receptor is composed of 2 chains, Tbp1 and Tbp2. H. influenzae is a
CC non-encapsulated or non-typable bacterium responsible for a wide range
CC of human diseases. Iron is an essential nutrient for the growth of these
CC bacteria, and they can utilise human transferrin as a source of iron.
CC Antibodies which block the access of the transferrin receptor to
CC its iron source prevent bacterial growth. The transferrin receptor, or
CC fragments, therefore, are good vaccine candidates. The full length Tbp2
CC protein is produced in low amounts in Escherichia coli. However, the
CC yield can be enhanced by truncation of the 3' end of the gene. An
CC immunogenic composition comprising (or encoding) the immunogenic
CC truncated analogue can be used to induce protection against a disease

CC caused by a bacterial pathogen that produces the transferrin receptor.
CC The immunogenic truncated analogue is also useful as an antigen in
CC immunoassays for the detection of Haemophilus transferrin receptor
CC antibodies, while the nucleic acid molecule can be used as a
CC hybridisation probe for the detection of other transferrin receptor
CC genes.
SQ Sequence 5033 BP; 1822 A; 837 C; 940 G; 1434 T;

Query Match 11.98; Score 331.2; DB 1; Length 5033;
Best Local Similarity 50.58; Pred. No. 4.7e-85;
Matches 1354; Conservative 0; Mismatches 1133; Indels 193; Gaps 15;

Qy 173 aaaaacgaagtaactggcttgggaaagtagtgaataattatcacgaaatgagtaaaaaac 232
Db 2355 ATAATGAAGTAACCTGACTTGGCAAAATATCAAAACTAGTGAAGTATCAGCCGAGAAC 2414
Qy 233 aaatcttggctatcgtagtatttaactcgctatgacctgtgatttcggtggtggaacaag 292
Db 2415 AAGTATTAAATATTCGTGATCTAACAGCTATGATCCAGGGATTTCAAGTTAGAACAAAG 2474
Qy 293 gtccggtgcaagtagtggctatccattcgaggtgtagataaaaacccgtgtcagcttac 352
Db 2475 GTCGCGGTGCAAGTCTCGGATATTCTATTCGTGTATGGACAGAAATAGAGTGTCTTAT 2534
Qy 353 ttgttgatgggtaccacaagcgcacagttatcatcagctaggttc----- 398
Db 2535 TAGTAGATGGTTTACCTCAACAGCAATCTTATGTAGTGCAAGCCCTTTAGTGTCTGTT 2594
Qy 399 ----agatgtctaattgggtgcaataatgagattgagtagtgaataacattcgttcaattg 454
Db 2595 CAGGATATTCCTGGCACTGGTCAATTAATGAATTAATGAATGAATGAATGAATGAATGA 2654
Qy 455 agttaacgaaggaagttctgcggaatattgctctggtgcgcatggtggtgctattg 514
Db 2655 AAATAAGCAAGGGGGGAGTCTTCTGAGTATGTAATGGAGCACTAGTGTGTCTGTAA 2714
Qy 515 gtttctactcaagatgcgcagattattataaagaggggcagcattggggttagata 574
Db 2715 CATTTCAAGCAAAATCAGCAGCCGATATCTTAGAAGGAGACAAATCATGTTGGGAATTTCAA 2774
Qy 575 gtaagacctcttgcagcaaaaaatagccatttttacagctctatcgagcggctggtg 634
Db 2775 CTAAAAATGCTTATTCAAGCAAAAAATAAGGCTTTACCCATTCTTTAGCTGTAGCAGAA 2834
Qy 635 aggcgggtgttttgaagcacttgttgcactaccacacacgcggttaaaagagaccacaa 694
Db 2835 AACAAAGTGGATTGGAAGGGCTAGCCATTTACACTCAACGAAATTCATTTGAAACCCCAAG 2894
Qy 695 ttcatccgagggcaataataataaaacataataatcggcggtataacggcgtttgaaatac 754
Db 2895 TCCATAAAGATGCATTAAGGGCGTACAAGGTTAAGTTCGATTAATCGCC-----ACAACA 2950
Qy 755 gctacgactttaccacaaattccgcacagaaatgctccctggaggatctcctttaaattg 814
Db 2951 GATAAATCTTCAGGATACTTGTGATACAAGGTGAGTG-----TCCAATGGTG 2999
Qy 815 aagatacttcccaacacattagattgtactcctcgtgcaagggttaagtgaacgcgcata 874
Db 3000 ATGACAAAGTGTGCAGC----- 3015
Qy 875 atttcccagtgagaacatttccgggaataatagcctgaagcgcacaaacagcttgagcaga 934
Db 3016 -----CAAGCCACTCGCACTT 3032
Qy 935 ttcttatcgactgagcagctctcagcccaagaataatccggttaaagatcgcatgacac 994
Db 3033 TATCCACCCCAAGGAAACCGTRAGCGTTTCAGATTATACGGGGGCTTAACCGTATCAAC 3092
Qy 995 caaaccttttagattacaagagtaattctgttttttgaagtttgggtatcaacttcaact 1054
Db 3093 CTAATCAATGAATATGAAAGCCAGTCTTGTTTAAAGAGGAGGATATCATTTTTCTG 3152

Qy 815 agataccttgcccaacattagattgtactcctcgtcgaagggttaagttgaaccgcgata 874
Dy 2970 ATGACAAGTGTGCAGC-----2985
Qy 875 atttccagtgagaacatttcgcggaataatacgcctcgaagagcgcaaacagcgttgagcaga 934
Dy 2986 -----CAAACCACTGCANAAT 3002
Qy 935 ttccttatgcactgagcagctctcagcccaagaataataccggtaaaagatgcattgcac 994
Dy 3003 TATCCCCCAAGCGAACCCTTAACGGTTTACAGATTATACGGGGCTTAACCGTATCAAC 3062
Qy 995 caaaccttttagattacaagagtaattctgttttttatgaagtttgctatcaactcaact 1054
Dy 3063 CTAATCCCAATGAATATGAAGCCAGCTGTGGTGTGTTTAAAGAGGGGTATCATTTTCTG 3122
Qy 1055 gctctcattatcttcgcgaactcttagaagatacaaaaacgcgtacgatatcogtgata 1114
Dy 3123 AACACACTATATGTGTGTTTGTGATTTTGAATTCACACAAACAAATTTGATATCCGTGATA 3182
Qy 1115 tgcaacgcagcttactatacaaaagacgattataacttatcacttaggaactatgttt 1174
Dy 3183 TGACATTTCCCGCTTATTAAGATCAACAGAAAACGGGATGATAGAACTGGCCCTTTT 3242
Qy 1175 atgaagggttaataatttttagatgcttagtgttcaagccaaggtaccccttatgggttgc 1234
Dy 3243 ATCCAAAGCAAGATTATGTGTCATATCAACGTTATTGAGGATGGCCGA-----GGCGTTA 3296
Qy 1235 gctatagcattggaagttttttgatgaagcgtcaccacaacagcgtgttaggattcaact 1294
Dy 3297 ATATGAAGTGGGGCTTATTTTCGATGAACACCAATAGAAAACAGCGTGTAGTATTGAAT 3356
Qy 1295 ataaataaaacagagaaataatcgtggttggatagcattaaac---tcagtgccggata 1351
Dy 3357 ATATTTACGAAAATGAAGAACAAAGCGGCATCAITGACAAAGCAGTGTAAAGTGTAAATC 3416
Qy 1352 aacaagatttgaactatagaccgggtcacatcgcttgcattgtgagcattatcctctggtg 1411
Dy 3417 AACAAACATCATCTTGACAGTTATATGCGACATPACGCAITTGCACTTTATCCTAATC 3476
Qy 1412 tagataaaattcgcgcgcactttgggataaaattcgttctatgtatgcagaactagcgata 1471
Dy 3477 CAAGTAAGAAATGCGCCCGACCTTGATAAACCTTATTCATCTACTACTGTTCTGTATAGAA 3536
Qy 1472 ataataccagaagaaagcatcgtgctcattcatttagaatttgataaagcgtctaaatgctg 1531
Dy 3537 ATGTTTAAAGAAAACATATATGTTGCAATTTGAATTTAGAGAAAATAATCAACAAA 3596
Qy 1532 gtcaaggcgtatttaaccaaaccacaactgaatttaggg-----ttgggctttg 1582
Dy 3597 ATTGCTTACTCATCAAAATTTGCTTCAATCTGGTTTGTGACTTTTACTTCAGCGCTTC 3656
Qy 1583 atcgatttaactcgtttagatcgtatggggtatagctgcccataataccaaagcggtt 1642
Dy 3657 AGCATAAAGATTATTAATCTCGACGTGTACCGCTACGGCAAAATATTATTTTCAGGGACAG 3716
Qy 1643 ataccagctacgcggtagagggcgttttagataatcctatatatttatcgcgcgcatccac 1702
Dy 3717 TTGCTGGTAAACGAAGAATGGTTACGAAAACAACTTACTTATCTACFCAAAACCAAAAG 3776
Qy 1703 gcagatttgaacggtatctctgttgtaataatacacgcgcg-----gacatcttaaac 1756
Dy 3777 TAGATTTTGTAGGACAAATCATTTGATATATTAAGAGTATGCTCTCTTAATTACAGCGACT 3836
Qy 1757 gtgaacgcgttaaaattaaagcgatagccattttgttagcttcgc---gatctagtga 1813
Dy 3837 GTAAAGTCGGTTAATTAAGGAAAAATATTAATTTCGCAGCAGCGCAATAATATGGCAT 3896
Qy 1814 taagcgatgtggtatttgattagggggtgctttttgatcaacatcgatttaaatctg 1873
Dy 3897 TAGGAAATACATGATTAGGTTTAGTATTTCGTATGACGTATCTCGTACAAAAGCTA 3956

Qy 1874 atgatcgtgacacattagccgaacttatcgaaaatttgcttggtaattggtggattacgc 1933
Dy 3957 ATGAATCAACATATTAGTGTGTTGTAATAATTTAAATTTCTCTTTGGAATACTGGTATTGCA 4016
Qy 1934 ttaaaccaacagaagtttgatcgcgttttcttatcgcatttcaaacggtttttagagtgctg 1993
Dy 4017 TAAACCAACAGGAATGGCTTGATCTTTCTTATCGCTTTCTTACTGGATTATGAATCCTA 4076
Qy 1994 catctatgaactttatggtaaacgctgatactatttgggcttaaaagataaacgaatatgac 2053
Dy 4077 GTTTTGTGAATGATGTGTTGGGGTATGTTGCGCAATAATAGCGATGTTTATGTAGTA 4136
Qy 2054 aacgcgcgaacgttagccaccagtttagagccagaaaaatcgactaatcatgatattggag 2113
Dy 4137 AA-----TTTAAGCCTGAAACATCTCGTAACCAAGATTGGTC 4175
Qy 2114 ttacgttaaaaggtcaatttggttaccttgatgtgagcgtatttccgttaataaactataaaa 2173
Dy 4176 TCGTCTAAAGGGGATTTTGGTAATATTGAGATCAGTCATTTTAGTAATGCTTATCGAA 4235
Qy 2174 atatga-----ttgcgacagcatgtaaaagaataatacaaaaaatcacactgttctata 2227
Dy 4236 ATCTTATCGCCTTTGCTGAAGAACTTAGTAAAAATGGAACTCTGGAAGGGCAATTATG 4295
Qy 2228 actaccataattcaagatgtagcactaaacgggataaaattttagctgctaaatttgact 2287
Dy 4296 GATATCAATATGACAAAATGCAAAATTAGTTGGCGTAAATATAAATTAACGCGCAATTAGATT 4355
Qy 2288 tacacgtattttatctatgctgcagatggttttttattcatcagttgcttataaacogtg 2347
Dy 4356 TTAATGTTTTATGAAAACGTATTCCCTACGGTTGTTGATGCAACATTTTGTCTTAAACCGAG 4415
Qy 2348 taaaagtaaaagagcgaaactaacccagactcaagactcgcatacgtagcgtaaacgatcctattc 2407
Dy 4416 TAAAAGTTAAAGATC---AAAAATCAATGCTGTTGGCTCCGTAGACGAGTTATTAT 4472
Qy 2408 tagatgcgattcagccagcacgctatgtcttgattcgcgctacgatacccccagaagaaa 2467
Dy 4473 TTGATGCCATTACGCCAGCGCTTATATCATTTGTTTAGCTATGATCATCCAAGTAATA 4532
Qy 2468 aatggggaattggcattactaccacctattttaaagccaaaaacgcgcgatgagtgccag 2527
Dy 4533 CTTCGGGAATTAATACAAATGTTTACTCAATCAAAAGCAAAATCTCAAAATGAATTGCTAG 4592
Qy 2528 gcacacgtcatccagcgtatcatgc-----gttgatttagtggaacacgtaccggtt 2580
Dy 4593 GACAACGTGCTATTCGGTAAACAAATTCAGGAATGTAAATCAACAGAAATTTACTCGGG 4652
Qy 2581 ctgtgtacacccatgattaccgg-----ttacatcaattataaaactcacaccttagctg 2637
Dy 4653 CATGSCATATCTTAGATGTATCGGGTTATTACATGGCGAATAAANAATATTATGCTTCGAT 4712
Qy 2638 gaggaaattataatgtgactaaatcgtaaatatttccatttgggaatcagtgccgaatccg 2697
Dy 4713 TAGGGATATAATTAATTAATCAACTATCGCTATGTTACTTGGGAAGCGGTGCGTCAACAG 4772
Qy 2698 gtggaatgcagtaaaccaagacc---ggggtagcaattacacatcgtatttgccgctcgg 2754
Dy 4773 CACAAGTGGCGGTCAATCAACATCAAAATGTTGTTAGCTATCTACTCGCTACGACATCAG 4832
Qy 2755 ggagaaatttcagtttagcatttgaagtgtta 2791
Dy 4833 GACGAACATATACCTTAACATTAGAAATGAATTTCTA 4869

RESULT 14

Q94444

ID Q94444 standard; DNA; 5009 BP.

AC Q94444;

DT 09-OCT-1996 (first entry)

DE Bacterial transferrin receptor operon (H. influenzae strain Minna).

KW Tbp1; Tbp2; transferrin receptor operon; vaccine; antigen;

KW non-typable strain; Haemophilus influenzae; meningitis; ss.

[illegible]

```
Qy 1751 taaactgtgaacccgctgaataataaagcgatagccatttggtagcttcgcg---gac 1807
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3819 GAGACTGTAAAGTGGGTGTAATTAAGAGGAAAAATTAATTTCCAGCAGCAATAATA 3878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1808 tagtgataagcagatgtatgtgatttgggatttaggggtcggtttgatcaaatgattta 1867
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3879 TGGCATTAGGGAATAACGTTGATTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAG 3938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1868 aatcgtgatccggtggacacttagcgaacttatcgaataatgctctgggaatggtgga 1927
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3939 AAGCTAATGAATCAACTATTAGTGTGTGAATTAATTAATTAATTAATTAATTAAT 3998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1928 ttacgcttaaaccaacagattgttatcgcttcttcttcttcgatttcaaacggttttagag 1987
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3999 TTGTCATAAAACCAACGGAATGGCTTGATCTTCTTATCGCTTCTTACTGGATTAGAA 4058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1988 tgcgtcattctatgaactttatggttaaacgtgatcatcatattgggcttaaaagataaacgaat 2047
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4059 ATCCTAGTATTTTCTGAAATGTATGTTGGCGGTATGGTGGCAAGAAATGACGAGGTTATG 4118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2048 atgtgaacgcgcgaacgttagccaccagttagagccagagaaaatcgactaatcatgaga 2107
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4119 TAGGTAAA-----TTTAGCCTGAACATCTCGTAACCAAGAGT 4157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2108 ttggagtttagcttaaaagttcaattgtgtaccttggatgtgagctatttccgtaataact 2167
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4158 TTGGTCTCGCTCTAAAGCGGATTTGCTAATATTGAGATCAGTCATTTTAGTAATGCTT 4217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2168 ataaaaatatgattgagcagcatgtaaa---agaataatacaaaaaatcacactgtttct 2224
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4218 ATCGAAATCTTATCGCCCTTGTGTAAGAACTTAGTAAAAATGGAATGGAAGGCAATT 4277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2225 ataacaccataattcaaatgtagcactaaacgggataaatttagtcgttaaatg 2284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4278 ATGGATATCAATATCCAAAATGCAAAATAGTTGGCGTAATAATACTGCACAAATAG 4337
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2285 acttacacggttttttctatgctgcagatggttttttattcattcagttcttataacc 2344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4338 ATTTTAATGGTTTATGGAAACGTATTCCTACGGTGTGTATGCAACATTTGCTTATAACC 4397
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2345 gtgtaaaagttaaaagcggaaacacacgactcaagactcgatagcgtaaacatcccta 2404
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4398 AAGTAAAAAGTTAAAGATC---AAAAATCAATGCTGTGTTAGCCTCCGTAAGCAGTTATT 4454
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2405 ttctagatgcattcagcagcagcagctatgctgttgattcggtatcgatcacccagaag 2464
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4455 TATTGTATGCCATTACGCCCGCCGCTTATATCATTTGTTAGGCTATGATCATCCAACTA 4514
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2465 aaaaatggggaattggcattactaccacttatcttaagcgaacaaacgcgcgatgaggtgg 2524
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4515 ATACTTGGGAATTAATCAATGTTTACTCAATCAAAAGCAAAATCTCAAAATGAATTGC 4574
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2525 caggcacacgtcaatcagcgatacatcgc-----gttgattaggtggcaaacgacgcg 2577
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4575 TAGGAAAAAGCGTATAGTAACTAATTAAGGGATGTAATAATCAACAAGAAAACTTACTC 4634
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2578 gtctctgtacacccatataattaccgg---ttaacataattataaaactacaccttac 2634
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4635 GGGCATGGCATATCTTAGATGATCGGGTATTACATGGGAATAAAAAATATTATGCTTTC 4694
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2635 gtggagggaattataatgtgactaatcgtlaaatattccacttgggaatcagtgqcccaat 2694
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4695 GATTAGGATATATAATTTTCACTATCGCTATGTTACTTGGGAAGCGGTGCGTCAAA 4754
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2695 cogggtgtaatgcagtaaaccaagaccg---gggtagcaattacactcgatttggcgctc 2751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4755 CAGCAACAAGGTGCGTCAATCAACATCAAAATGTTGGTAGCTATATCGCTACGCAGCAT 4814
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2752 cggggagaaaatttcagtttagcatttgaatgaagtttta 2791
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4815 CAGGAGGAACATATACCTTAACTTAGAAATGAATTTCTA 4854
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 15

```
Q94443
ID O94443 standard; DNA; 5033 BP.
AC O94443;1996 (first entry)
DT 09-OCT-1996
DE Bacterial transferrin receptor operon (H. influenzae strain Eagan).
KW Tbp1; Tbp2; transferrin receptor operon; vaccine; antigen;
KW non-typable strain; Haemophilus influenzae; meningitis; ss.
OS Haemophilus influenzae strain Eagan.
FH Key Location/Qualifiers
FT cds 169..2151
    /*tag= a
    /product= Tbp2
    2165..4904
    /*tag= b
    /product= Tbp1
FN W09513370-A1.
PD 18-MAY-1995.
PF 07-NOV-1994; CA0616.
PR 08-NOV-1993; US-148968.
PR 29-DEC-1993; US-175116.
PA (CONN-) CONNAUGHT LAB LTD.
PI Chong P, Gray-owen S, Harkness R, Klein M, Loosmore S;
PI Murdin A, Schryvers A, Yang Y;
PI WPI; 95-194089/25.
DR P-PSDB; R77885-86.
CC Nucleic acids encoding Haemophilus transferrin receptor - used to
CC develop prods for detection and in diagnosis, prevention and
CC treatment of Haemophilus infection.
CC Claim 12; Fig A-A; 231pp; English.
CC The present DNA shows the transferrin receptor (Tfr) operon consisting of
CC two genes (Tbp1 and Tbp2) arranged in tandem and which are transcribed
CC from a single promoter, from Haemophilus influenzae type b, strain Eagan.
CC H. influenzae Tfr is iron- and/or haemin-regulated and a putative fur-
CC binding site has been identified upstream of Tbp2. Antibodies blocking
CC this binding site may prevent bacterial growth. Fragments of the Tfr
CC (or its genes) are useful in vaccines to provide protection against, e.g.
CC bacterial meningitis. An advantage of using the Tfr is that it shares
CC homology with Tfr of other H. influenzae strains including non-typable
CC strains.
SQ Sequence 5033 BP; 1822 A; 840 C; 937 G; 1434 T;

Query Match 11.8%; Score 329.6; DB 1; Length 5033;
Best Local Similarity 50.5%; Pred. No. 1.4e-84;
Matches 135; Conservative 0; Mismatches 1134; Indels 193; Gaps 15;

Qy 173 aaacgaagtaactggcttgggaaagtagtgaataattatcacgaaatgagtaaaatc 232
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2355 ATATGAAGTAACGTGGACTTGGCAAAATATCAAACTAGTGAAGTATCAGCCGAGAAC 2414
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 233 aaatctctgtattcgtgatttaactcgtatgacctgtgtatttcggtgggaaacaag 292
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2415 AAGTATTAAATATTCGTGATCTAACACGCTATGATCCAGGGATTTTCAGTTGTAGAACAAG 2474
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 293 gtccgctgcaagtagtggctatccattcgaggttagataaaaacgcgtcagcttac 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2475 GTCCGGTGCANGTCTGGATATTCTATTCGTGTATGGACAGAAATAGATGCTCTTAT 2534
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 353 ttgttgatgggctaccacaagcgacaggttatcatcagctagggtc----- 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2535 TAGTAGATGGTTTACCCTCAACGCAATCTATGTAGTGAAGCCCTTTAGTTGCTCGTT 2594
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 399 ----agatgctaatgtgtgcaattaatgagattgagattgaaacaattcgttcaattg 454
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2595 CAGGATATTCGACATGGTGCATTAATGAATTAATGAATGAATGAATGAATGAATGAATGA 2654
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 455 agttaagcaaaaggaagaattctcgcgaataatgctctggtgcgcagtggtggtctattg 514
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2655 AAATAAGCAAGGGGGGAGTCTTCTTCTGATGTATGTAATGGAGACACTAGCTGTTCTGTA 2714
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 515 gtttcgtactaaagatgcgcaggtattattaaagggggcagcattggggcttagata 574
```

Db 2715 CATTTCAAAGCAATCAGCAGCCGATATCTTTAGAAGGAGACAAATCATCGGGAATTCAAA 2774
Qy 575 gtaagaccttatagcagcaaaaaatagccattttttacagctctatcgacggcgtggtg 634
Db 2775 CTAANAATGCTTATTTCAAGCAAAAATAAGGCTTTACCCATTTTAGCTGTAGCAGAA 2834
Qy 635 agcggtggttttgaagcactgtttatgcaactccgacacggttaagagaccacaaa 694
Db 2835 AACAAAGTGGGATTTGAAGGGTAGCCATTTACACTCAACGAANAATCAATTTGAAACCAAG 2894
Qy 695 ttcatctcgagcgaataaataataaataatttcggcgataaccggctttgaaatc 754
Db 2895 TCCATAAAGATGCATTAAGGCGGTACAAAGCTTATGATCGATTATCGCC-----ACRACA 2950
Qy 755 gctacgaccttaccocaaattccgcacagaatgctcctgggagatctcctttaaattggtg 814
Db 2951 GATAAATCTTCAGGATACTTTGTGATACAAGGTGAGTG-----TCCAAATGGTG 2999
Qy 815 aagatacttgcccaacattagatgtactcctcgtgcaagggttaagtgaaccgcgata 874
Db 3000 ATGCAAGTGTGCAGC----- 3015
Qy 875 atttcccgatgagaacatttccggaataatacgcctgaagagcgcaaacagcttgagcaga 934
Db 3016 -----CAAGCCACTGCGACTT 3032
Qy 935 ttccttatgcactgagcagctctcagcccaagaataataccggttaagatgcgattgcac 994
Db 3033 TATCACCAAGGAAACCCGTAAAGCGTTTCAGATTATACGGGGCTAACGGTATCAAAAC 3092
Qy 995 caaacctttagattacaagagtaattcgtttttatgaagtttggctatcaactcaact 1054
Db 3093 CTAATCCAAATGAATATGAAGCGCAGTCTGGTTTTTAAGAGGAGGGGTATCATTTTTCTG 3152
Qy 1055 cgtctcattactctggcgcaactcttagaagatacaaaaacgcgtacagatccgctgata 1114
Db 3153 AACAACTATTATTTGGTGTATTTTGNATTCACACACAAAATTTGATATCCGTGATA 3212
Qy 1115 tgcacacgcgcgtactatatacaaaagcagatataaacttcaacttaggaactatgttt 1174
Db 3213 TGACATTTCCCGCTATTATTAAGCCCAACAGAAAGCGGATGATAGTAGTGTCTCTTTT 3272
Qy 1175 atgaagggtataatttttagatggcttagtggttcagacccaagatcccttatgggtgc 1234
Db 3273 ATCCAAATCAAGATCATGGTGCAATATCAACATATTGAGGATGGCA-----GAGGCGTTA 3326
Qy 1235 gctatagccatgtgaagttttttgatgaacgctcaccaacacgcgttttaggttcaacct 1294
Db 3327 AATATGCAAGTGGGCTTTATTTTCGATGAACACCATAGANAACACGGGTATGTAAT 3386
Qy 1295 ataaataaaaccagagaaataatcgtcgttggtagatattaaactc---agtcgcgata 1351
Db 3387 ATATTTAGAAAATAAGAACAAAGCGGCATCATTTGACAAAAGCAGTGTAAAGTGTCTAATC 3446
Qy 1352 aacaagatttgaactatatagcggcgtacacgtctgcatgttagcgattatcctgtgg 1411
Db 3447 AACAAAACATCATCTTGACAGTTATATGCGACATACCGCATTTGCGTCTTATCCTAATC 3506
Qy 1412 tagataaaaattgcgcgcgcactttggataaattcttggctctatgtatgaactgagcgta 1471
Db 3507 CAAGTAAGAATTGCGGCCCAACACTTGATATAACCTTATTCATCTATCGTCTCTGATAGA 3566
Qy 1472 ataattaccaagaaaagcatcgtgtcatcttatttagaatttgataaa----- 1518
Db 3567 ATGTTTAAAGAAAAACATAATATGTTGCAATTTGAATTTAGAGAAAAAATAATCAACAAA 3626
Qy 1519 --gcgctaaatgctgggtcaagcgctatttaaccacaaacccacaaactgaattgggtgg 1576
Db 3627 ATTGGCTTACTCATCAAAATGCTTCAATCTTGTTTGTGACCTTTACTTCACGGCTTC 3686
Qy 1577 gctttgatcgatttaattcgtttatggatcatggggatgactgcccataataccaaaag 1636

Db 3687 AGCATAAAGATTTATTTAACTCGAGCTGTTATCGCTACGCGCAGATAGTATTCCAGGAAAC 3746
Qy 1637 gcggttataccagctaccgcggttagagggcgttttagataatccatataattatgcgcgcg 1696
Db 3747 CTGGTGAACCTGCTTAACCAAGAAATGTTTGAATCACAACCTTACTTATACCCAAAAC 3806
Qy 1697 atccacgcagattatgaacgggtatcttctgtgtaataatacacgcggc-----gacatct 1750
Db 3807 CAGAGCCATATTTTGCAGGACAGATCATTTGTAATATCAAGGTAGCTCTCTTAATTACA 3866
Qy 1751 taaactgtaaccgcgtataaattaaagcgatagccattttgttagcttccgc---gac 1807
Db 3867 GAGACTTAAAGTGCCTTAATTAAGGAAATAATTATTATTCGCACGCAATAATA 3926
Qy 1808 tagtgaataagcggtatgctggtattggttaggggtaggggtaggggttttgatacaaatcgatta 1867
Db 3927 TGGCATTAGGGAATAAGCTTGAATTTAGTATTAGGTATTCGGTATGAGTATCTCGTACAA 3986
Qy 1868 aatctgatgatccgtagacacttagccgaactatcgaaattggtcttgaatggtggga 1927
Db 3987 AAGCTAATGAATCAACTATTAGTGTGTTGTAATTTAAAAATTTCTTTGGAATACTGGTA 4046
Qy 1928 ttagccttaaaccaacagagtttggatcgcctttcttcatcgcatttcaaacgggttttagag 1987
Db 4047 TTGTCATAAAACCAACGGAATGCTTGATCTTCTTATCGCITTCTTACTGGATTAGAA 4106
Qy 1988 tgcctgactctatgaactttatggttaaacgctgataatattgggcttaagatgaacgaat 2047
Db 4107 ATCCTAGTTTTTCTGAATATGATGTTGGGGTATGTTGGCAAGAAATGACGAGGTTATG 4166
Qy 2048 atgtgcaacgcggcgaacagtagccaccagtttagccagacagaaaaactgaactaatcagga 2107
Db 4167 TAGTAAA-----TTTAAGCCTGAAACATCTCGTAACCAAGAGT 4205
Qy 2108 ttggagtagctttaaagggtcaatttggttacccttgatgagctatttccgttaataact 2167
Db 4206 TTGCTCTCGCTCTAAAAGGGGATTTTGGTAATATTGAGATCAGTCATTTTAGTAATGCTT 4265
Qy 2168 ataaaaataagttgcagacagcatgtaaa---agaaataacaaaaatcacactgtttct 2224
Db 4266 ATCGAAATCTTATCGCCTTTGCTGAAGAACTTAGTAAAAATGGAACCTGGAAGGCGCAAT 4325
Qy 2225 ataactaccataatattcaagatgtagcactaaacgggataaaatttagtgcgtctaaattg 2284
Db 4326 ATGGATATCATATATGCAAAAATGCAAAATAGTTGGCGGTAAATATTAAGTCACAATTAG 4385
Qy 2285 acttacacggtattttatctatgctgccagatggttttttattcatcagttgcttataacc 2344
Db 4386 ATTTTAATGGTTTATGCAAAAGGTATTCCTACGGTTGGTATGCAACATTTGCTTATRAAC 4445
Qy 2345 gtgtaaaagtaaaagagcggaaactaacccagctcaagactcgatagcgtaaacgactccta 2404
Db 4446 AAGTAAAGTTAAAGATC---AAAAATCAATGCTGTTTAGCCTCGTAAAGCAGTATT 4502
Qy 2405 ttctagatgctgattcagcagcagcgtatgcttgcattcggtcgttagcagcaccagaag 2464
Db 4503 TATTTGATCCCATTCAGCCCGCGTTATATCATGTTGTTTAGGCTATGATCATCCAGTA 4562
Qy 2465 aaaaatgggaaattggttactactaccacttattctaagcgaacaaacgcgcatgaggtgg 2524
Db 4563 ATACTTGGGAATTAATACAATGTTTACTCAATCAAAAGCAAAATCTCAAAATGAATTGC 4622
Qy 2525 caggcacacgtcatcagcgatacatcgc-----gttgatttaggtggcgaactgaccg 2577
Db 4623 TAGAAAACGTGCATTAGGTAACAAATTCAGGGGATGTAAAAATCAACAGAAATTTACTC 4682
Qy 2578 gttcttggtcacaccatgatattaccgg---tttacctcaaatataaaaaactcacaccttac 2634
Db 4683 GGGCATGGCATATCTTAGATGTATCGGGTTATTACATGGCGGAATAAAAAATTATGCTTC 4742
Qy 2635 gtggaggaaattataatgtgactaaatgtaaatattccacttgggaatcaggtgcgcgaat 2694
Db 4743 GATTAGGGATATATAATTTTATTTCAACTATGCTATGTTACTTGGGAAGCGGTGCTCAAA 4802

..

Thu Sep 14 09:51:05 2000

us-08-753-750a-1.rng

Page 25

Qy 2695 ccggtgtgaatgcagtaaaccaagaccg---gggtagcaattacactcgtattggcgctc 2751
Db 4803 CAGCACAAAGGTGCGGTCAATCAACATCAAAATGTTGGTAGCTATACTCGCTACGCAGCAT 4862
Qy 2752 cggggagaaatttcagtttagcatttgaagtgttta 2791
Db 4863 CAGGACGAAACTATACCTTACATTAGAAATGAAATTCTA 4902

Search completed: September 14, 2000, 00:18:18
Job time: 3871 sec

THIS PAGE BLANK (USPTO)